GP 1636

#4

IN THE U.S. PATENT AND TRADEMARK OFFICE

Applicant: BERNSTEIN, Jeanne et al

Application No.: 09/695,293 Group:

Filed: October 25, 2000 Examiner:

For: VARIANTS OF ALTERNATIVE SPLICING

L E T T E R

Honorable Commissioner of Patents
and Trademarks
Washington, D.C. 20231

March 8, 2001
2786-0140P

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55(a), the applicant hereby claims the right of priority based on the following application(s):

Country	Application No.	Filed
ISRAEL	132558	10/25/99

A certified copy of the above-noted application(s) is(are) attached hereto.

If necessary, the Commissioner is hereby authorized in this, concurrent, and future replies, to charge payment or credit any overpayment to Deposit Account No. 02-2448 for any additional fees required under 37 C.F.R. 1.16 or under 37 C.F.R. 1.17; particularly, extension of time fees.

Respectfully submitted,

BIRCH, STEWART, KOLASCH & BIRCH, LLP

By: _____
JOHN CASTELLANO
Reg. No. 35,094
P. O. Box 747
Falls Church, Virginia 22040-0747

Attachment
(703) 205-8000
/dp

Birch Stewart et al
703-205-8000
2786-0140 P
Jeanne Berstein et al
09/695,293

מדינת ישראל
STATE OF ISRAEL

Ministry of Justice
Patent Office

משרד המשפטים
לשכת הפטנטים

This is to certify that annexed hereto is a true copy of the documents as originally deposited with the patent application particulars of which are specified on the first page of the annex.

זאת לתעודה כי רצופים בזה העתקים נכונים של המסמכים שהופקדו לכתחילה עם הבקשה לפטנט לפי הפרטים הרשומים בעמוד הראשון של הנספח.

This 31-01-2001 היום

רשם הפטנטים
Commissioner of Patents

נתאשר
Certified

חוק הפטנטים, תשכ"ז – 1967
PATENTS LAW, 5727-1967

בקשה לפטנט
Application For Patent

לשימוש הלשכה
For Office Use

מספר: Number	132558
תאריך: Date	25 -10- 1993
הוקדם/נדחה: Ante/Post-dated	

אני, (שם המבקש, מענו ולגבי גוף מאוגד: מקום התאגדותו)
I, (Name and address of applicant, and in case of body corporate-place of incorporation)

קומפיוגן בע"מ, חברה ישראלית מרחוב פנחס רוזן 72, תל אביב 69512, ישראל
Compugen Ltd., Israeli Company of 72, Pinchas Rozen St., Tel Aviv 69512, Israel

בעל אמצאה מכח: הדין — Right of Law — שמה הוא
Owner, by virtue of — of an invention the title of which is

(בעברית) (Hebrew)
וריאנטים של חיתוך-חיבור אלטרנטיבי

(באנגלית) (English)
Variants of alternative splicing

מבקש בזאת כי ינתן לי עליה פטנט
Hereby apply for a patent to be granted to me in respect thereof.

* בקשת חלוקה Application of Division	* בקשת פטנט מוסף Appl. for Patent of Addition	דרישת דין קדימה* Priority Claim		
מבקשת פטנט from application	* לבקשה/לפטנט to Patent/Appl.	מספר/סימן Number/Mark	תאריך Date	מדינת האיגוד Convention Country
No. מס'	No. מס'			
Dated מיום	Dated מיום			
P.o.A.: General כללי * יפוי כח: filed in case P128131 הוגש בעניין				

המען למסירת מסמכים בישראל
Address for Service in Israel

REINHOLD COHN AND PARTNERS
Patent Attorneys
P.O.B. 4060, Tel-Aviv
C. 120583

חתימת המבקש
Signature of Applicant

For the Applicants,
REINHOLD COHN AND PARTNERS
By : —

This — היום 24 בחודש October שנת 1999 of Year

לשימוש הלשכה
For Office Use

טופס זה כשהוא מוטבע בחותם לשכת הפטנטים ומושלם במספר ובתאריך ההגשה, הנו אישור להגשת הבקשה שפרטיה רשומים לעיל.
This form, impressed with the Seal of the Patent Office and indicating the number and date of filing, certifies the filing of the application the particulars of which are set out above.

* מחק את המיותר
Delete whatever is inapplicable

וריאנטים של חיתוך-חיבור אלטרנטיבי

Variants of alternative splicing

Compugen Ltd.

קומפיוגן בע"מ

C. 120583

FIELD OF THE INVENTION

The present invention concerns novel nucleic acid sequences, vectors and host cells containing them, amino acid sequences encoded by said sequences, and antibodies reactive with said amino acid sequences, as well as pharmaceutical compositions comprising any of the above. The present invention further concerns methods for screening for candidate activators or deactivators utilizing said amino acid sequences.

BACKGROUND OF THE INVENTION

Alternative splicing (AS) is an important regulatory mechanism in higher eukaryotes (P.A. Sharp, *Cell* 77, 805-8152 (1994). It is thought to be one of the most important mechanisms for differential expression related to tissue or development stage specificity. It is known to play a major role in numerous biological systems, including human antibody responses, and sex determination in *Drosophila*, (S. Stamm, M.Q. Zhang, T.G. Marr and D.M. Helfman, *Nucleic Acids Research* **22**, 1515-1526 (1994); B. Chabot, *Trends Genet.* **12**, 472-478 (1996); R.E. Breitbart, A. Andreadis, B. Nadal-Ginard, *Annual Rev. Biochem.*, **56**, 467-495 (1987); C.W. Smith, J.G. Patton, B. Nadal-Ginard, *Annu. Rev. Genet.*, **27**, 527-577 (1989)).

Until recently it was commonly believed that alternative splicing existed in only a small fraction of genes (about 5%). A recent observation based on literature survey of known genes revises this conservative estimate to as high as an estimate that at least 30% of human genes are alternatively spliced (M.S. Gelfand, I. Dubchak, I. Draluk and M. Zorn, *Nucleic Acids Research* **27**, 301-302 (1999). The importance of the actual frequency of this phenomenon lies not only in the direct impact on the number of proteins created (100,000 human genes, for example, would be translated to a much higher number of proteins), but also in the diversity of functionality derived from the process.

Several mechanisms at different stages may be held responsible for the complexity of higher eukaryote which include: alternative splicing at the transcription level, RNA editing at the post-transcriptional level, and post-translational modifications are the ones characterized to date.

GLOSSARY

In the following description and claims use will be made, at times, with a variety of terms, and the meaning of such terms as they should be construed in accordance with the invention is as follows:

"Variant nucleic acid sequence" – the sequence shown in any one of SEQ ID NO: 1 to SEQ ID NO: 17, sequences having at least *90% identity* (see below) to said sequence and *fragments* (see below) of the above sequences of least 20 b.p. long. These sequences are sequences coding for a novel, naturally occurring, alternative splice variants of native and known genes. It should be emphasized that the novel variants of the present invention are naturally occurring sequences resulting from alternative splicing of genes and not merely truncated, mutated or fragmented forms of known sequences which are artificially produced.

"Variant product – also referred at times as the "variant protein" or ***"variant polypeptide"*** – is an amino acid sequence encoded by the variant nucleic acid sequence which is a naturally occurring mRNA sequence obtained as a result of alternative splicing. The amino acid sequence may be a peptide, a protein, as well as peptides or proteins having *chemically modified* amino acids (see below) such as a glycopeptide or glycoprotein. The variant products are shown in any one of SEQ ID NO: 18 to SEQ ID NO: 34. The term also includes *homologies* (see below) of said sequences in which one or more amino acids has been added, deleted, *substituted* (see below) or *chemically modified* (see below) as well as *fragments* (see below) of this sequence having at least 10 amino acids.

"Nucleic acid sequence" – a sequence composed of DNA nucleotides, RNA nucleotides or a combination of both types and may includes natural nucleotides, chemically modified nucleotides and synthetic nucleotides.

"Amino acid sequence" – a sequence composed of any one of the 20 naturally appearing amino acids, amino acids which have been *chemically modified* (see below), or composed of synthetic amino acids.

"Fragment of variant nucleic acid sequence" – novel short stretch of nucleic acid sequences of at least 20 b.p., which does not appear as a continuous stretch in the *original nucleic acid sequence* (see below). The fragment may be a sequence which was previously undescribed in the context of the published RNA and which affects the amino acid sequence encoded by the known gene. For example, where the variant nucleic includes a sequence which was not included in the original sequence (for example a sequence which was an intron in the original sequence) the fragment may contain said additional sequence. The fragment may also be a region which is not an intron, which was not present in the original sequence. For example where the variant lacks a non-terminal region which was present in the original sequence. The two stretches of nucleotides spanning this region (upstream and downstream) are brought together by splicing in the variant, but are spaced from each by the spliced out region in the original sequence and are thus not continuous in the original sequence. A continuous stretch of nucleic acids comprising said two sparing stretches of nucleotides is not present in the original sequence and thus falls under the definition of fragment.

"Fragments of variant products" - novel amino acid sequences coded by the *"fragment of variant nucleic acid sequence"* defined above.

"Homologues of variants" – amino acid sequences of variants in which one or more amino acids has been added, deleted or replaced. The addition, deletion or replacement should be in the regions or adjacent to regions where the variant differs from the *original sequence* (see below).

"Conservative substitution" - refers to the substitution of an amino acid in one class by an amino acid of the same class, where a class is defined by common physicochemical amino acid side chain properties and high substitution frequencies in homologous proteins found in nature, as determined, for example, by a standard Dayhoff frequency exchange matrix or BLOSUM matrix. [Six general classes of amino acid side chains have been categorized and include: Class I (Cys); Class II (Ser, Thr, Pro, Ala, Gly); Class III (Asn, Asp, Gln, Glu); Class IV (His, Arg, Lys); Class V (Ile, Leu, Val, Met); and Class VI (Phe, Tyr, Trp). For example, substitution of an Asp for another class III residue such as Asn, Gln, or Glu, is a conservative substitution.

"Non-conservative substitution" - refers to the substitution of an amino acid in one class with an amino acid from another class; for example, substitution of an Ala, a class II residue, with a class III residue such as Asp, Asn, Glu, or Gln.

"Chemically modified" - when referring to the product of the invention, means a product (protein) where at least one of its amino acid resides is modified either by natural processes, such as processing or other post-translational modifications, or by chemical modification techniques which are well known in the art. Among the numerous known modifications typical, but not exclusive examples include: acetylation, acylation, amidation, ADP-ribosylation, glycosylation, GPI anchor formation, covalent attachment of a lipid or lipid derivative, methylation, myristlyation, pegylation, prenylation, phosphorylation, ubiqutination, or any similar process.

"Biologically active" - refers to the variant product having some sort of biological activity, for example, some physiologically measurable effect on target cells, molecules or tissues.

"Immunologically active" defines the capability of a natural, recombinant or synthetic varient product, or any fragment thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies. Thus, for example, an immunologically active fragment of variant product denotes a fragment which retains some or all of the immunological properties of the variant product, e.g can bind specific anti-variant product antibodies or which can elicit an immune response which will generate such antibodies or cause proliferation of specific immune cells which produce variant.

"Optimal alignment" - is defined as an alignment giving the highest percent identity score. Such alignment can be performed using a variety of commercially available sequence analysis programs, such as the local alignment program LALIGN using a ktup of 1, default parameters and the default PAM. A preferred alignment is the one performed using the CLUSTAL-W program from MacVector (TM), operated with an open gap penalty of 10.0, an extended gap penalty of 0.1, and a BLOSUM similarity matrix. If a gap needs to be inserted into a first sequence to optimally align it with a second sequence, the percent identity is calculated using only the residues that are paired with a corresponding amino acid residue (i.e., the calculation does not consider residues in the second sequences that are in the *"gap"* of the first sequence). In case of alignments of known gene sequences with that of the new variant, the optimal alignment invariably included aligning the identical parts of both sequences together, then keeping apart and unaligned the sections of the sequences that differ one from the other.

"Having at least 90% identity" - with respect to two amino acid or nucleic acid sequence sequences, refers to the percentage of residues that are identical in the two sequences when the sequences are optimally aligned. Thus, 90% amino acid sequence identity means that 90% of the amino acids in two or more optimally aligned polypeptide sequences are identical, however this definition explicitly excludes sequences which are 100% identical with the original sequence from which the variant of the invention was varied.

"Isolated nucleic acid molecule having an variant nucleic acid sequence" - is a nucleic acid molecule that includes the coding variant nucleic acid sequence. Said isolated nucleic acid molecule may include the variant nucleic acid sequence as an independent insert; may include the variant nucleic acid sequence fused to an additional coding sequences, encoding together a fusion protein in which the variant coding sequence is the dominant coding sequence (for example, the additional coding sequence may code for a signal peptide); the variant nucleic acid sequence may be in combination with non-coding sequences, e.g., introns or control elements, such as promoter and terminator elements or 5' and/or 3' untranslated regions, effective for expression of the coding sequence in a suitable host; or may be a vector in which the variant protein coding sequence is a heterologous.

"Expression vector" - refers to vectors that have the ability to incorporate and express heterologous DNA fragments in a foreign cell. Many prokaryotic and eukaryotic expression vectors are known and/or commercially available. Selection of appropriate expression vectors is within the knowledge of those having skill in the art.

"Deletion" - is a change in either nucleotide or amino acid sequence in which one or more nucleotides or amino acid residues, respectively, are absent.

"Insertion" or "addition" - is that change in a nucleotide or amino acid sequence which has resulted in the addition of one or more nucleotides or amino acid residues, respectively, as compared to the naturally occurring sequence.

"Substitution" - replacement of one or more nucleotides or amino acids by different nucleotides or amino acids, respectively. As regards amino acid sequences the substitution may be conservative or non- conservative.

"Antibody" – refers to IgG, IgM, IgD, IgA, or IgG antibody. The definition includes polyclonal antibodies or monoclonal antibodies. This term refers to whole antibodies or fragments of the antibodies comprising the antigen-binding domain of the anti-variant product antibodies, e.g. antibodies without the Fc portion, single chain antibodies, fragments consisting of essentially only the variable, antigen-binding domain of the antibody, etc.

Distinguishing antibody" – an antibody capable of binding to the variant product and not the original amino acid sequence from which it has been varied, or an antibody capable of binding to the original nucleic acid sequence and not to the variant production.

"Activator" - as used herein, refers to a molecule which mimics the effect of the natural variant product or at times even increases or prolongs the duration of the biological activity of said product, as compared to that induced by the natural product. The mechanism may be by any mechanism known to prolonging activities of biological molecules such as binding to receptors; prolonging the lifetime of the molecules; increasing the activity of the molecules on its target; increasing the affinity of molecules to its receptor; inhibiting degradation or proteolysis of the molecules, or mimicking the biological activity of the variants on their targets, etc. Activators may be polypeptides, nucleic acids,

carbohydrates, lipids, or derivatives thereof, or any other molecules which can bind to and activate the variant product.

"Deactivator" or ***("Inhibitor")*** - refers to a molecule which modulates the activity of the variant product in an opposite manner to that of the activator, by decreasing or shortening the duration of the biological activity of the variant product. This may be done by any mechanism known to deactivate or inhibit biological molecules such as block of the receptor, block of active site, competition on binding site in target, enhancement of degradation, etc. Deactivators may be polypeptides, nucleic acids, carbohydrates, lipids, or derivatives thereof, or any other molecules which bind to and modulate the activity of said product.

"Treating a disease" - refers to administering a therapeutic substance effective to ameliorate symptoms associated with a disease, to lessen the severity or cure the disease, or to prevent the disease from occurring.

"***Detection***" – refers to a method of detection of a disease, disorder, pathological or normal condition. This term may refer to detection of a predisposition to a disease as well as for establishing the prognosis of the patient by determining the severity of the disease.

"Probe" – the variant nucleic acid sequence, or a sequence complementary therewith, when used to detect presence of other similar sequences in a sample. The detection is carried out by identification of hybridization complexes between the probe and the assayed sequence. The probe may be attached to a solid support or to a detectable label.

"Original sequence" – the amino acid or nucleic acid sequence from which the variant of the invention have been varied as a result of alternative slicing.

SUMMARY OF THE INVENTION

The present invention is based on the finding of several novel, naturally occurring splice variants, which are naturally occurring sequences obtained by alternative splicing of known genes. The novel splice variants of the invention are not merely truncated forms, fragments or mutations of known genes, but rather novel sequences which naturally occur within the body of individuals.

The term *"alternative splicing"* in the context of the present invention and claims refers to: intron inclusion, exon exclusion, addition or deletion of terminal sequences in the variant as compared to the original sequences, as well as to the possibility of *"intron retention"*. Intron retention is an intermediate stage in the processing of RNA transcripts, where prior to production of fully processed mRNA the intron (naturally spliced in the original sequence) is retained in the variant. These intermediately processed RNAs may have physiological significance and are also within the scope of the invention.

The novel variant products of the invention may have the same physiological activity as the original peptide from which they have been varied (although perhaps at a different level); may have an opposite physiological activity from the activity featured by the original peptide from which they are varied; may have a completely different, unrelated activity to the activity of the original from which they are varied; or alternatively may have no activity at all and this may lead to various diseases or pathological conditions.

The novel variants may also serve for detection purposes, i.e. their presence or level may be indicative of a disease, disorder, pathological or normal condition or alternatively the ratio between the level variants and the level original peptide from which they were varied, or the ratio to other variants may be indicative to a disease, disorder, pathological or normal condition.

For example, for detectional purposes, it is possible to establish differential expression of various variants in various tissues. A certain variant may be expressed mainly in one tissue, while the original sequence from which it has been varied, or another variant may, be expressed mainly in another tissue.

Understanding of the distribution of the variants in various tissues may be helpful in basic research, for understanding the physiological function of the genes as well as may help in targeting pharmaceuticals or developing pharmaceuticals.

The study of the variants may also be helpful to distinguish various stages in the life cycles of the same type of cells which may also be helpful for development of pharmaceuticals for various pathological conditions in which cell cycles is non-normal, notably cancer.

Thus the detection may by determination of the presence or the level of expression of the variant within a specific cell population, comprising said presence or level between various cell types in a tissue, between different tissues and between individuals.

Thus the present invention provides by its first aspect, a novel isolated nucleic acid molecule comprising or consisting of any one of the coding sequence SEQ ID NO: 1 to SEQ ID NO: 17, fragments of said coding sequence having at least 20 nucleic acids (provided that said fragments are continuous stretches of nucleotides not present in the original sequence from which the variant was varied), or a molecule comprising a sequence having at least 90%, identity to SEQ ID NO: 1 to SEQ ID NO: 17, provided that the molecule is not completely identical to the original sequence from which the variant was varied.

The present invention further provides a protein or polypeptide comprising or consisting of an amino acid sequence encoded by any of the above nucleic acid sequences, termed herein *"variant product"*, for example, an amino acid sequence having the sequence as depicted in any one of SEQ ID NO: 18 to SEQ ID NO: 34, fragments of the above amino acid sequence having a length of at least 10 amino acids coded by the above fragments of the nucleic acid sequences, as well as homologues of the above amino acid sequences in which one or more of the amino acid residues has been substituted (by conservative or non-conservative substitution) added, deleted, or chemically modified.

The deletions, insertions and modifications should be in regions, or adjacent to regions, wherein the variant differs from the original sequence.

For example, where the variant is different from the original sequence by addition of a short stretch of 10 amino acids, in the terminal or non- terminal portion of the peptide, the invention also concerns homologues of that variant where the additional short stretch is altered for example, it includes only 8 additional amino acids, includes 13 additional amino acids, or it includes 10 additional amino acids, however some of them being conservative or non-conservative substitutes of the original additional 10 amino acids of the novel variants. In all cases the changes in the homolog, as compared to the original sequence, are in the same regions where the variant differs from the original sequence, or in regions adjacent to said region.

Another example is where the variant lacks a non-terminal region (for example of 20 amino acids) which is present in the original sequence (due for example to exon exclusion). The homologues may lack in the same region only 17 amino acids or 23 amino acids. Again the deletion is in the same region where the variant lacks a sequence as compared to the original sequence, or in a region adjacent thereto.

It should be appreciated that once a man versed in the art's attention is directed to the importance of a specific region, due to the fact that this region differs in the variant as compared to the original sequence, there is no problem in derivating said specific region by addition to it, deleting from it, or substituting some amino acids in it. Thus homologues of variants which are derivated from the variant by changes (deletion, addition, substitution) only in said region as well as in regions adjacent to it are also a part of the present invention. Generally, if the variant is distinguished from the original sequence by some sort of physiological activity, then the homolog is distinguished from the original sequence in essentially the same manner.

The present invention further provides nucleic acid molecule comprising or consisting of a sequence which encodes the above amino acid sequences, (including the fragments and homologues of the amino acid sequences). Due to the degenerative nature of the genetic code, a plurality of alternative nucleic acid

sequences, beyond those depicted in any one of SEQ ID NO: 1 to SEQ ID NO: 17, can code for the amino acid sequence of the invention. Those alternative nucleic acid sequences which code for the same amino acid sequences codes by the sequence SEQ ID NO: 1 to SEQ ID NO: 17 are also an aspect of the of the present invention.

The present invention further provides expression vectors and cloning vectors comprising any of the above nucleic acid sequences, as well as host cells transfected by said vectors.

The present invention still further provides pharmaceutical compositions comprising, as an active ingredient, said nucleic acid molecules, said expression vectors, or said protein or polypeptide.

These pharmaceutical compositions are suitable for the treatment of diseases and pathological conditions, which can be ameliorated or cured by raising the level of any one of the variant products of the invention.

By a second aspect, the present invention provides a nucleic acid molecule comprising or consisting of a non-coding sequence which is complementary to that of any one of SEQ ID NO: 1 to SEQ ID NO: 17, or complementary to a sequence having at least 90% identity to said sequence (with the proviso added above) or a fragment of said two sequences (according to the above definition of fragment). The complementary sequence may be a DNA sequence which hybridizes with any one of SEQ of ID NO: 1 to SEQ ID NO: 17 or hybridizes to a portion of that sequence having a length sufficient to inhibit the transcription of the complementary sequence. The complementary sequence may be a DNA sequence which can be transcribed into an mRNA being an antisense to the mRNA transcribed from any one of SEQ ID NO: 1 to SEQ ID NO: 17 or into an mRNA which is an antisense to a fragment of the mRNA transcribed from any one of SEQ ID NO: 1 to SEQ ID NO: 17 which has a length sufficient to hybridize with the mRNA transcribed from SEQ ID NO: 1 to SEQ ID NO: 17, so as to inhibit its translation. The complementary sequence may also be the mRNA or the fragment of the mRNA itself.

The nucleic acids of the second aspect of the invention may be used for therapeutic or diagnostic applications for example as probes used for the detection of the variants of the invention. The presence of the variant transcript or the level of the variant transcript may be indicative of a multitude of diseases, disorders and various pathological as well as normal conditions. In addition or alternatively, the ratio of the level of the transcripts of the variants of the invention may also be compared to that of the transcripts of the original sequences from which have been varied, or to the level of transcript of other variants, and said ratio may be indicative to a multitude of diseases, disorders and various pathological and normal conditions.

The present invention also provides expression vectors comprising any one of the above defined complementary nucleic acid sequences and host cells transfected with said nucleic acid sequences or vectors, being complementary to those specified in the first aspect of the invention.

The invention also provides anti-variant product antibodies, namely antibodies directed against the variant product which specifically bind to said variant product. Said antibodies are useful both for diagnostic and therapeutic purposes. For example said antibody may be as an active ingredient in a pharmaceutical composition as will be explained below.

The present invention also provides pharmaceutical compositions comprising, as an active ingredient, the nucleic acid molecules which comprise or consist of said complementary sequences, or of a vector comprising said complementary sequences. The pharmaceutical composition thus provides pharmaceutical compositions comprising, as an active ingredient, said anti-variant product antibodies.

The pharmaceutical compositions comprising said anti-variant product antibodies or the nucleic acid molecule comprising said complementary sequence, are suitable for the treatment of diseases and pathological conditions where a therapeutically beneficial effect may be achieved by neutralizing the variant (either at the transcript or product level) or decreasing the amount of the variant product or

blocking its binding to its target, for example, by the neutralizing effect of the antibodies, or by the effect of the antisense mRNA in decreasing the expression level of the variant sequence.

According to the third aspect of the invention the present invention provides methods for detecting the level of the transcript (mRNA) of said variant product in a body fluid sample, or in a specific tissue sample, for example by use of probes comprising or consisting of said coding sequences; as well as methods for detecting levels of expression of said product in tissue, e.g. by the use of antibodies capable of specifically reacting with the variant products of the invention. Detection of the level of the expression of the variant of the invention in particular as compared to that of the original sequence from which it was varied or compared to other variant sequences all varied from the same original sequence may be indicative of a plurality of physiological or pathological conditions.

The method, according to this latter aspect, for detection of a nucleic acid sequence which encodes the variant product in a biological sample, comprises the steps of:

(a) providing a probe comprising at least one of the nucleic acid sequences defined above;

(b) contacting the biological sample with said probe under conditions allowing hybridization of nucleic acid sequences thereby enabling formation of hybridization complexes;

(c) detecting hybridization complexes, wherein the presence of the complexes indicates the presence of nucleic acid sequence encoding the variant product in the biological sample.

The method as described above is qualitative, i.e. indicates whether the transcript is present in or absent from the sample. The method can also be quantitative, by determining the level of hybridization complexes and then calibrating said levels to determining levels of transcripts of the desired variant in the sample.

Both qualitative and quantitative determination methods can be used for diagnostic, prognostic and therapy planning purposes.

By a preferred embodiment the probe is part of a nucleic acid chip used for detection purposes, i.e. the probe is a part of an array of probes each present in a known location on a solid support.

The nucleic acid sequence used in the above method may be a DNA sequence an RNA sequence, etc; it may be a coding or a sequence or a sequence complementary thereto (for respective detection of RNA transcripts or coding-DNA sequences). By quantization of the level of hybridization complexes and calibrating the quantified results it is possible also to detect the level of the transcript in the sample.

Methods for detecting mutations in the region coding for the variant product are also provided, which may be methods carried-out in a binary fashion, namely merely detecting whether there is any mismatches between the normal variant nucleic acid sequence of the invention and the one present in the sample, or carried-out by specifically detecting the nature and location of the mutation.

The present invention also concerns a method for detecting variant product in a biological sample, comprising the steps of:

(a) contacting with said biological sample the antibody of the invention, thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

wherein the presence of said antibody-antigen complex correlates with the presence of variant product in said biological sample.

Many diseases are diagnosed by detecting the presence of antibodies against a protein characterizing the disease in the blood, serum or any other body fluid of the patient. The present invention also concerns a method for detecting anti-variant antibody in a biological sample, comprising:

(a) contacting said sample with the variant product of the invention, thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

wherein the presence of said antibody-antigen complex correlates with the presence of anti-variant antibody in the sample.

As indicated above, both methods (for detection of variant product and for detection of the anti-variant antibody) can be quantitized to determine the level or the amount of the variant or antibody in the sample, alone or in comparison to the level of the original amino acid sequence from which it was varied or compared to the level of antibodies against the original amino acid sequence, and qualitative and quantitative results may be used for diagnostic, prognostic and therapy planning purposes.

The invention also concerns distinguishing antibodies, i.e. antibodies capable of binding either to the variant product or to the original sequence from which the variant has been varied, while not binding to the original sequence or the variant product respectively. These distinguishing antibodies may be used for detection purposes.

By yet another aspect the invention also provides a method for identifying candidate compounds capable of binding to the variant product and modulating its activity (being either activators or deactivators). The method includes:

(i) providing a protein or polypeptide comprising an amino acid sequence substantially as depicted in any one of SEQ ID NO: 18 to 34, or a fragment of such a sequence;

(ii) contacting a candidate compound with said amino acid sequence;

(iii) measuring the physiological effect of said candidate compound on the activity of the amino acid sequences and selecting those compounds which show a significant effect on said physiological activity.

The present invention also concerns compounds identified by the above methods described above, which compound may either be an activator of the variant product or a deactivator thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

In order to understand the invention and to see how it may be carried out in practice, a preferred embodiment will now be described, by way of non-limiting example only, with reference to the accompanying drawings, in which:

Fig. 1 is a comparison between the amino acid sequence of SEQ ID NO: 18 and the original sequence from which it has been varied;

Fig. 2 is a comparison between the amino acid sequence of SEQ ID NO: 19 and the original sequence from which it has been varied;

Fig. 3 is a comparison between the amino acid sequence of SEQ ID NO: 20 and the original sequence from which it has been varied;

Fig. 4 is a comparison between the amino acid sequence of SEQ ID NO: 21 and the original sequence from which it has been varied;

Fig. 5 is a comparison between the amino acid sequence of SEQ ID NO: 22 and the original sequence from which it has been varied;

Fig. 6 is a comparison between the amino acid sequence of SEQ ID NO: 23 and the original sequence from which it has been varied;

Fig. 7 is a comparison between the amino acid sequence of SEQ ID NO: 24 and the original sequence from which it has been varied;

Fig. 8 is a comparison between the amino acid sequence of SEQ ID NO: 25 and the original sequence from which it has been varied;

Fig. 9 is a comparison between the amino acid sequence of SEQ ID NO: 26 and the original sequence from which it has been varied;

Fig. 10 is a comparison between the amino acid sequence of SEQ ID NO: 27 and the original sequence from which it has been varied;

Fig. 11 is a comparison between the amino acid sequence of SEQ ID NO: 28 and the original sequence from which it has been varied;

Fig. 12 is a comparison between the amino acid sequence of SEQ ID NO: 29 and the original sequence from which it has been varied;

Fig. 13 is a comparison between the amino acid sequence of SEQ ID NO: 30 and the original sequence from which it has been varied;

Fig. 14 is a comparison between the amino acid sequence of SEQ ID NO: 31 and the original sequence from which it has been varied;

Fig. 15 is a comparison between the amino acid sequence of SEQ ID NO: 32 and the original sequence from which it has been varied;

Fig. 16 is a comparison between the amino acid sequence of SEQ ID NO: 33 and the original sequence from which it has been varied;

Fig. 17 is a comparison between the amino acid sequence of SEQ ID NO: 34 and the original sequence from which it has been varied.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Example I: Comparison of variants with original sequences

Original sequences were obtained from GenBank Version 110. Comparison between the original sequences and the noval variant sequences was made using the Pileup application from the GCG suite version 10.0 (January 1999), with the defalut values:

Gap creation penalty (GapWeight): 8

Gap extension penalty (GapLengthWeight): 2

The comparison is shown in Fig. 1 to 17 which show the comparison of each of the variant products depicted in SEQ ID NO: 18 to 34 with the original sequence from which it was varied.

The following is a table that compares the sequences of the variants of the invention to the original sequences from which they were varied and indicates ehere the variant differs from the original sequence. The terminology NV-1 to NV-17 corresponds to SEQ ID NO: 1 to SEQ ID NO:17.

TABLE

Accession #	New Variant #	Description of the new variant
BPI_HUMAN	NV-1	Replacement of the last C-terminal 92 aa of the protein with alternative 100 aa, which has an additional potential transmembrane domain.
BAL_HUMAN	NV-2	Replacement of 121 C-terminal amino acids of the original protein by alternative 7 amino acids.
FA12_HUMAN	NV-3	Insertion of 30 amino acids in the catalytic domain of the original protein.
TIM1_HUMAN	NV-4	Replacement of 52 C-terminal amino acids of the origginal protein by alternative 9 amino acids. Missing 3 disulfide bonds.
TIM1_HUMAN	NV-5	Replacement of 94 C-terminal amino acids of the original protein by alternative 10 amino acids. Has only 1 out of 5 diS bonds, has an extra Cys in the added 10 amino acids.
TIM1_HUMAN	NV-6	Replacement of one of the exons of the original protein by a homologous exon.
TIM1_HUMAN	NV-7	Deletion of 47 amino acids of the original protein. Missing 2 disulfide bonds.
TIM1_HUMAN	NV-8	Deletion of 10 amino acids of the original protein. Missing one out of 2 glycosylation sites.
UROK_HUMAN	NV-9	Insertion of 62 amino acids in long A chain.
PAI1_HUMAN	NV-10	Replacement of 140 C-terminal amino acids of the original protein, including the active site and 2 out of 3 glycosylation sites, by alternative 12 amino acids.
PAI1_HUMAN	NV-11	Replacement of 17 C-terminal amino acids of the original protein by alternative 14 amino acids.
CTGF_HUMAN	NV-12	Deletion of 32 amino acids, at positions 152-185 of the original protein.
DCC_HUMAN	NV-13	Replacement of 1118 C-terminal amino acids of the original protein, including cytoplasmic domain, transmembrane region and part of the extracellular domain, by alternative 12 amino acoids. The deleted region contains all 5 fibronectin type domains and 1 out of the 4 Ig-like C2 type domain.
DCC_HUMAN	NV-14	Replacement of 28 C-terminal amino acids from the cytoplasmic domain of the original protein by alternative 13 amino acids.
MGR4_HUMAN	NV-15	Deletion of 58 amino acids from the extracellular domain of the original protein.
MCR_HUMAN	NV-16	Replacement of 147 C-terminal amino acids, including part of the steroid binding domain of the original protein, by alternative 8 amino acids.
MCR_HUMAN	NV-17	Deletion of 147 C-terminal amino acids, including part of the steroid binding domain of the original protein.

Example II: Designation of the original sequences

Each novel variant of the invention is varied from an original sequence which has a known designation. The designation of the RNA sequences of the original sequences from which it was varied and the Accession Number of the original sequence are given below. First, information concerning the original sequence is given and then designation of the novel variants of the invention is given as NV-1 to NV-17 corresponding to SEQ ID NO: 1 to SEQ ID NO: 17.

BPI_HUMAN

BACTERICIDAL PERMEABILITY-INCREASING PROTEIN

FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.
SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR LEUKOCYTES (PMN) GRANULES.
TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES.
DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N- AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE INGESTED.

NV-1:

BAL_HUMAN

BILE-SALT-ACTIVATED LIPASE (BAL)
CHOLESTEROL ESTERASE

FUNCTION: CATALYZES FAT AND VITAMIN ABSORPTION. ACTS IN CONCERT WITH PANCREATIC LIPASE AND COLIPASE FOR THE COMPLETE DIGESTION OF DIETARY TRIGLYCERIDES.
CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL + A FATTY ACID ANION.
CATALYTIC ACTIVITY: A STERYL ESTER + H(2)O = A STEROL + A FATTY ACID.
ENZYME REGULATION: ACTIVATED BY BILE SALTS CONTAINING A 7-HYDROXYL GROUP.
TISSUE SPECIFICITY: MAMMARY GLAND, AND PANCREAS.
SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.

NV-2:

FA12_HUMAN

COAGULATION FACTOR XII

FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE GENERATION OF BRADYKININ AND ANGIOTENSIN.
CATALYTIC ACTIVITY: CLEAVES SELECTIVELY ARG-|-ILE BONDS AND ACTIVATES COAGULATION FACTORS VII AND XI.
PTM: O- AND N-GLYCOSYLATED.
DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR XIIA ACTIVATES FACTOR XI TO FACTOR XIA.

NV-3:

TIM1_HUMAN

METALLOPROTEINASE INHIBITOR 1

FUNCTION: COMPLEXES WITH METALLOPROTEINASES (SUCH AS COLLAGENASES) AND IRREVERSIBLY INACTIVATE THEM. ALSO MEDIATES ERYTHROPOIESIS IN VITRO; BUT, UNLIKE IL-3, IT IS SPECIES-SPECIFIC, STIMULATING THE GROWTH AND DIFFERENTIATION OF ONLY HUMAN AND MURINE ERYTHROID PROGENITORS.
PTM: THE ACTIVITY OF TIMP-1 IS DEPENDENT ON THE PRESENCE OF DISULFIDE BONDS.
SIMILARITY: BELONGS TO THE TIMP FAMILY.

NV-4:

TIM1_HUMAN

METALLOPROTEINASE INHIBITOR 1

NV-5:

TIM1_HUMAN

METALLOPROTEINASE INHIBITOR 1

NV-6:

TIM1_HUMAN

METALLOPROTEINASE INHIBITOR 1

NV-7

TIM1_HUMAN

METALLOPROTEINASE INHIBITOR 1

NV-8

UROK_HUMAN

UROKINASE-TYPE PLASMINOGEN ACTIVATOR

FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR THERAPY OF THROMBOLYTIC DISORDERS.
CATALYTIC ACTIVITY: SPECIFIC CLEAVAGE OF ARG-|-VAL BOND IN PLASMINOGEN TO FORM PLASMIN.
SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
PHARMACEUTICAL: AVAILABLE UNDER THE NAME ABBOKINASE (ABBOTT). USED IN PULMONARY EMBOLISM (PE) TO INITIATES FIBRINOLYSIS.

NV-9:

PAI1_HUMAN

PLASMINOGEN ACTIVATOR INHIBITOR-1, ENDOTHELIAL

FUNCTION: THIS INHIBITOR ACTS AS "BAIT" FOR TISSUE PLASMINOGEN ACTIVATOR, UROKINASE, AND PROTEIN C. ITS RAPID INTERACTION WITH TPA MAY FUNCTION AS A MAJOR CONTROL POINT IN THE REGULATION OF FIBRINOLYSIS.
DISEASE: HIGH CONCENTRATIONS OF THIS PROTEIN HAVE BEEN ASSOCIATED WITH HUMAN THROMBOEMBOLIC DISEASE.
MISCELLANEOUS: PAI1 IS INACTIVATED BY PROTEOLYTIC ATTACK OF THE UROKINASE-TYPE (U-PA) AND THE TISSUE-TYPE (TPA), CLEAVING THE 369(R)-370(M) BOND.

NV-10:

PAI1_HUMAN

PLASMINOGEN ACTIVATOR INHIBITOR-1, ENDOTHELIAL

NV-11:

CTGF_HUMAN

CONNECTIVE TISSUE GROWTH FACTOR

FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
SUBUNIT: MONOMER.
ALTERNATIVE PRODUCTS: A SHORTER FORM MAY BE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE: VARSPLIC 172 198 MISSING (IN SHORT FORM).
SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV

NV-12:

DCC_HUMAN

TUMOR SUPPRESSOR PROTEIN DCC

FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.
DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR POINT MUTATIONS MAY CAUSE BOTH

LYMPHATIC AND HEMATOGENOUS METASTASIS OF OESOPHAGEAL SQUAMOUS CELL CARCINOMAS.

SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.

NV-13:

DCC_HUMAN

TUMOR SUPPRESSOR PROTEIN DCC

NV-14:

MGR4_HUMAN

METABOTROPIC CLUTAMATE RECEPTOR 4

FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
TISSUE SPECIFICITY: STRONGLY EXPRESSED IN THE CEREBELLUM. EXPRESSED AT LOW LEVELS IN HIPPOCAMPUS, HYPOTHALAMUS AND THALAMUS. NO EXPRESSION DETECTED IN LIVER.
SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS. STRONGEST, TO MGLUR6.

NV-15:

MCR_HUMAN

MINERALOCORTICOID RECEPTOR

FUNCTION: RECEPTOR FOR BOTH MINERALOCORTICOIDS (MC) SUCH AS ALDOSTERONE AND GLUCOCORTICOIDS (GC) SUCH AS

CORTICOSTERONE OR CORTISOL. THE EFFECT OF MC IS TO INCREASE ION AND WATER TRANSPORT AND THUS RAISE EXTRACELLULAR FLUID VOLUME AND BLOOD PRESSURE AND LOWER POTASSIUM LEVELS.
SUBCELLULAR LOCATION: NUCLEAR.
DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY. NR3 SUBFAMILY.

NV-16:

MCR_HUMAN

MINERALOCORTICOID RECEPTOR

NV-17:

Example III: Variant nucleic acid sequence

The nucleic acid sequences of the invention include nucleic acid sequences which encode variant product and fragments and analogs thereof. The nucleic acid sequences may alternatively be sequences complementary to the above coding sequence, or to a region of said coding sequence. The length of the complementary sequence is sufficient to avoid the expression of the coding sequence. The nucleic acid sequences may be in the form of RNA or in the form of DNA, and include messenger RNA, synthetic RNA and DNA, cDNA, and genomic DNA. The DNA may be double-stranded or single-stranded, and if single-stranded may be the coding strand or the non-coding (anti-sense, complementary) strand. The nucleic acid sequences may also both include dNTPs, rNTPs as well as non naturally occurring sequences. The sequence may also be a part of a hybrid between an amino acid sequence and a nucleic acid sequence.

In a general embodiment, the nucleic acid sequence has at least 90%, identity with any one of the sequence identified as SEQ ID NO: 1 to SEQ ID NO: 17 provided that this sequence is not completely identical with that of the original sequence.

The nucleic acid sequences may include the coding sequence by itself. By another alternative the coding region may be in combination with additional coding sequences, such as those coding for fusion protein or signal peptides, in combination with non-coding sequences, such as introns and control elements, promoter and terminator elements or 5' and/or 3' untranslated regions, effective for expression of the coding sequence in a suitable host, and/or in a vector or host environment in which the variant nucleic acid sequence is introduced as a heterologous sequence.

The nucleic acid sequences of the present invention may also have the product coding sequence fused in-frame to a marker sequence which allows for purification of the variant product. The marker sequence may be, for example, a hexahistidine tag to provide for purification of the mature polypeptide fused to the marker in the case of a bacterial host, or, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., *et al. Cell* **37**:767 (1984)).

Also included in the scope of the invention are fragments as defined above also referred to herein as oligonucleotides, typically having at least 20 bases, preferably 20-30 bases corresponding to a region of the coding-sequence nucleic acid sequence. The fragments may be used as probes, primers, and when complementary also as antisense agents, and the like, according to known methods.

As indicated above, the nucleic acid sequence may be substantially a depicted in any one of SEQ ID NO: 1 to SEQ ID NO: 17 or fragments thereof or sequences having at least 90% identity to the above sequence as explained above. Alternatively, due to the degenerative nature of the genetic code, the sequence

may be a sequence coding for any one of the amino acid sequence of SEQ ID NO: 18 to SEQ ID NO: 34, or fragments or analogs of said amino acid sequence.

A. Preparation of nucleic acid sequences

The nucleic acid sequences may be obtained by screening cDNA libraries using oligonucleotide probes which can hybridize to or PCR-amplify nucleic acid sequences which encode the variant products disclosed above. cDNA libraries prepared from a variety of tissues are commercially available and procedures for screening and isolating cDNA clones are well-known to those of skill in the art. Such techniques are described in, for example, Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual (2nd Edition), Cold Spring Harbor Press, Plainview, N.Y. and Ausubel FM et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

The nucleic acid sequences may be extended to obtain upstream and downstream sequences such as promoters, regulatory elements, and 5' and 3' untranslated regions (UTRs). Extension of the available transcript sequence may be performed by numerous methods known to those of skill in the art, such as PCR or primer extension (Sambrook *et al., supra*), or by the RACE method using, for example, the Marathon RACE kit (Clontech, Cat. # K1802-1).

Alternatively, the technique of *"restriction-site"* PCR (Gobinda *et al. PCR Methods Applic.* 2:318-22, (1993)), which uses universal primers to retrieve flanking sequence adjacent a known locus, may be employed. First, genomic DNA is amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR can be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. *et al.*, *Nucleic Acids Res.* 16:8186, (1988)). The primers may be designed using OLIGO(R) 4.06 Primer Analysis

Software (1992; National Biosciences Inc, Plymouth, Minn.), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Capture PCR (Lagerstrom, M. *et al., PCR Methods Applic.* **1**:111-19, (1991)) is a method for PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA. Capture PCR also requires multiple restriction enzyme digestions and ligations to place an engineered double-stranded sequence into a flanking part of the DNA molecule before PCR.

Another method which may be used to retrieve flanking sequences is that of Parker, J.D., *et al., Nucleic Acids Res.*, **19**:3055-60, (1991)). Additionally, one can use PCR, nested primers and PromoterFinder™ libraries to *"walk in"* genomic DNA (PromoterFinder™; Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions. Preferred libraries for screening for full length cDNAs are ones that have been size-selected to include larger cDNAs. Also, random primed libraries are preferred in that they will contain more sequences which contain the 5' and upstream regions of genes.

A randomly primed library may be particularly useful if an oligo d(T) library does not yield a full-length cDNA. Genomic libraries are useful for extension into the 5' nontranslated regulatory region.

The nucleic acid sequences and oligonucleotides of the invention can also be prepared by solid-phase methods, according to known synthetic methods. Typically, fragments of up to about 100 bases are individually synthesized, then joined to form continuous sequences up to several hundred-bases.

B. Use of variant nucleic acid sequence for the production of variant products

In accordance with the present invention, nucleic acid sequences specified above may be used as recombinant DNA molecules that direct the expression of variant products.

As will be understood by those of skill in the art, it may be advantageous to produce variant product-encoding nucleotide sequences possessing codons other than those which appear in any one of SEQ ID NO: 1 to SEQ ID NO: 17 which are those which naturally occur in the human genome. Codons preferred by a particular prokaryotic or eukaryotic host (Murray, E. *et al. Nuc Acids Res.*, **17**:477-508, (1989)) can be selected, for example, to increase the rate of variant product expression or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, than transcripts produced from naturally occurring sequence.

The nucleic acid sequences of the present invention can be engineered in order to alter a variant product coding sequence for a variety of reasons, including but not limited to, alterations which modify the cloning, processing and/or expression of the product. For example, alterations may be introduced using techniques which are well known in the art, e.g., site-directed mutagenesis, to insert new restriction sites, to alter glycosylation patterns, to change codon preference, etc.

The present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a nucleic acid sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. Appropriate cloning and

expression vectors for use with prokaryotic and eukaryotic hosts are also described in Sambrook, *et al., (supra)*.

The present invention also relates to host cells which are genetically engineered with vectors of the invention, and the production of the product of the invention by recombinant techniques. Host cells are genetically engineered (i.e., transduced, transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the expression of the variant nucleic acid sequence. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to those skilled in the art.

The nucleic acid sequences of the present invention may be included in any one of a variety of expression vectors for expressing a product. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host. The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and related sub-cloning procedures are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate transcription control sequence (promoter) to direct mRNA synthesis. Examples of such promoters include: LTR or SV40 promoter, the *E.coli lac* or *trp* promoter, the phage lambda *PL* promoter, and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses.

The expression vector also contains a ribosome binding site for translation initiation, and a transcription terminator. The vector may also include appropriate sequences for amplifying expression. In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E.coli.*

The vector containing the appropriate DNA sequence as described above, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein. Examples of appropriate expression hosts include: bacterial cells, such as *E.coli, Streptomyces, Salmonella typhimurium*; fungal cells, such as yeast; insect cells such as *Drosophila* and *Spodoptera* Sf9; animal cells such as CHO, COS, HEK 293 or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein. The invention is not limited by the host cells employed.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the variant product. For example, when large quantities of variant product are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be desirable. Such vectors include, but are not limited to, multifunctional *E.coli* cloning and expression vectors such as *Bluescript*(R) (Stratagene), in which the variant polypeptide coding sequence may be ligated into the vector in-frame with sequences for the amino-terminal Met and the subsequent 7 residues of beta-galactosidase so that a hybrid protein is produced; *pIN* vectors (Van Heeke & Schuster *J. Biol. Chem.* **264**:5503-5509, (1989)); *pET* vectors (Novagen, Madison WI); and the like.

In the yeast *Saccharomyces cerevisiae* a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and

PGH may be used. For reviews, see Ausubel *et al. (supra)* and Grant *et al.*, (*Methods in Enzymology* **153**:516-544, (1987)).

In cases where plant expression vectors are used, the expression of a sequence encoding variant product may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of *CaMV* (Brisson *et al., Nature* **310**:511-514. (1984)) may be used alone or in combination with the omega leader sequence from TMV (Takamatsu *et al., EMBO J.*, **6**:307-311, (1987)). Alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi *et al., EMBO J.* **3**:1671-1680, (1984); Broglie *et al., Science* **224**:838-843, (1984)); or heat shock promoters (Winter J and Sinibaldi R.M., *Results Probl. Cell Differ.*, **17**:85-105, (1991)) may be used. These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. For reviews of such techniques, see Hobbs S. or Murry L.E. (1992) in McGraw Hill Yearbook of Science and Technology, McGraw Hill, New York, N.Y., pp 191-196; or Weissbach and Weissbach (1988) *Methods for Plant Molecular Biology*, Academic Press, New York, N.Y., pp 421-463.

Variant product may also be expressed in an insect system. In one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The variant product coding sequence may be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of variant coding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein coat. The recombinant viruses are then used to infect *S. frugiperda* cells or *Trichoplusia* larvae in which variant protein is expressed (Smith *et al., J. Virol.* 46:584, (1983); Engelhard, E.K. *et al., Proc. Nat. Acad. Sci.* **91**:3224-7, (1994)).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, a variant product coding sequence may be ligated into an adenovirus

transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome will result in a viable virus capable of expressing variant protein in infected host cells (Logan and Shenk, *Proc. Natl. Acad. Sci.* **81**:3655-59, (1984). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be required for efficient translation of a variant product coding sequence. These signals include the ATG initiation codon and adjacent sequences. In cases where variant product coding sequence, its initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous transcriptional control signals including the ATG initiation codon must be provided. Furthermore, the initiation codon must be in the correct reading frame to ensure transcription of the entire insert. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate to the cell system in use (Scharf, D. *et al.*, (1994) *Results Probl. Cell Differ.*, **20**:125-62, (1994); Bittner et al., *Methods in Enzymol* **153**:516-544, (1987)).

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., and Battey, I. (1986) Basic Methods in Molecular Biology). Cell-free translation systems can also be employed to produce polypeptides using RNAs derived from the DNA constructs of the present invention.

A host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the protein include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Post-translational processing which cleaves a *"pre-pro"* form of the protein may also be important for correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, 293, WI38, etc. have specific cellular machinery and characteristic mechanisms for such post-translational activities and may be chosen to ensure the correct modification and processing of the introduced, foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express variant product may be transformed using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clumps of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler M., *et al.*, *Cell* **11**:223-32, (1977)) and adenine phosphoribosyltransferase (Lowy I., *et al.*, *Cell* **22**:817-23, (1980)) genes which can be employed in *tk-* or *aprt-* cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, *dhfr* which confers resistance to methotrexate (Wigler M., *et al., Proc. Natl. Acad. Sci.* 77:3567-70, (1980)); *npt*, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. *et al., J. Mol. Biol.*, **150**:1-14, (1981)) and *als* or *pat*, which confer resistance to chlorsulfuron and

phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, *trpB*, which allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman S.C. and R.C. Mulligan, *Proc. Natl. Acad. Sci.* **85**:8047-51, (1988)). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate, GUS, and luciferase and its substrates, luciferin and ATP, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. *et. al.*, *Methods Mol. Biol.*, 55:121-131, (1995)).

Host cells transformed with a nucleotide sequence encoding variant product may be cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The product produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing nucleic acid sequences encoding variant product can be designed with signal sequences which direct secretion of variant product through a prokaryotic or eukaryotic cell membrane.

The variant product may also be expressed as a recombinant protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle, Wash.). The inclusion of a protease-cleavable polypeptide linker sequence between the purification domain and variant product is useful to facilitate purification. One such expression vector provides for expression of a fusion protein compromising a variant polypeptide fused to a polyhistidine region separated by an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC

(immobilized metal ion affinity chromatography, as described in Porath, *et al.*, *Protein Expression and Purification,* 3:263-281, (1992)) while the enterokinase cleavage site provides a means for isolating variant polypeptide from the fusion protein. *pGEX* vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to ligand-agarose beads (e.g., glutathione-agarose in the case of GST-fusions) followed by elution in the presence of free ligand.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, or other methods, which are well know to those skilled in the art.

The variant products can be recovered and purified from recombinant cell cultures by any of a number of methods well known in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

C. Diagnostic applications utilizing nucleic acid sequences

The nucleic acid sequences of the present invention may be used for a variety of diagnostic purposes. The nucleic acid sequences may be used to detect

and quantitate expression of the variant in patient's cells, e.g. biopsied tissues, by detecting the presence of mRNA coding for variant product. Alternatively, the assay may be used to detect soluble variant in the serum or blood. This assay typically involves obtaining total mRNA from the tissue or serum and contacting the mRNA with a nucleic acid probe. The probe is a nucleic acid molecule of at least 20 nucleotides, preferably 20-30 nucleotides, capable of specifically hybridizing with a sequence included within the sequence of a nucleic acid molecule encoding variant product under hybridizing conditions, detecting the presence of mRNA hybridized to the probe, and thereby detecting the expression of variant. This assay can be used to distinguish between absence, presence, and excess expression of variant product and to monitor levels of variant expression during therapeutic intervention. In addition, the assay may be used to compare the levels of the variant of the invention to the levels of the original sequence from which it has been varied or to levels of other variants, which comparison may have some physiological meaning.

The invention also contemplates the use of the nucleic acid sequences as a diagnostic for diseases resulting from inherited defective variant sequences, or diseases in which the ratio of the amount of the original sequence from which the variant was varied to the novel variants of the invention is altered. These sequences can be detected by comparing the sequences of the defective (i.e., mutant) variant coding region with that of a normal coding region. Association of the sequence coding for mutant variant product with abnormal variant product activity may be verified. In addition, sequences encoding mutant variant products can be inserted into a suitable vector for expression in a functional assay system (e.g., colorimetric assay, complementation experiments in a variant protein deficient strain of HEK293 cells) as yet another means to verify or identify mutations. Once mutant genes have been identified, one can then screen populations of interest for carriers of the mutant gene.

Individuals carrying mutations in the nucleic acid sequence of the present invention may be detected at the DNA level by a variety of techniques. Nucleic

acids used for diagnosis may be obtained from a patient's cells, including but not limited to such as from blood, urine, saliva, placenta, tissue biopsy and autopsy material. Genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki, *et al.*, *Nature* **324**:163-166, (1986)) prior to analysis. RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid of the present invention can be used to identify and analyze mutations in the gene of the present invention. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype.

Point mutations can be identified by hybridizing amplified DNA to radiolabeled RNA of the invention or alternatively, radiolabeled antisense DNA sequences of the invention. Sequence changes at specific locations may also be revealed by nuclease protection assays, such RNase and S1 protection or the chemical cleavage method (e.g. Cotton, *et alProc. Natl. Acad. Sci.* USA, **85**:4397-4401, (1985)), or by differences in melting temperatures. *"Molecular beacons"* (Kostrikis L.G. *et al.*, Science **279**:1228-1229, (1998)), hairpin-shaped, single-stranded synthetic oligo- nucleotides containing probe sequences which are complementary to the nucleic acid of the present invention, may also be used to detect point mutations or other sequence changes as well as monitor expression levels of variant product. Such diagnostics would be particularly useful for prenatal testing.

Another method for detecting mutations uses two DNA probes which are designed to hybridize to adjacent regions of a target, with abutting bases, where the region of known or suspected mutation(s) is at or near the abutting bases. The two probes may be joined at the abutting bases, e.g., in the presence of a ligase enzyme, but only if both probes are correctly base paired in the region of probe junction. The presence or absence of mutations is then detectable by the presence or absence of ligated probe.

Also suitable for detecting mutations in the variant product coding sequence are oligonucleotide array methods based on sequencing by

hybridization (SBH), as described, for example, in U.S. Patent No. 5,547,839. In a typical method, the DNA target analyte is hybridized with an array of oligonucleotides formed on a microchip. The sequence of the target can then be "read" from the pattern of target binding to the array.

D. Gene mapping utilizing nucleic acid sequences

The nucleic acid sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 20-30 bp) from the variant cDNA. Computer analysis of the 3' untranslated region is used to rapidly select primers that do not span more than one exon in the genomic DNA, which would complicate the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids or using instead radiation hybrids are rapid procedures for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60 bases. For a review of this technique, see Verma *et al., Human Chromosomes: a Manual of Basic Techniques*, (1988) Pergamon Press, New York.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in the OMIM database (Center for Medical Genetics, Johns Hopkins University, Baltimore, MD and National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD). The OMIM gene map presents the cytogenetic map location of disease genes and other expressed genes. The OMIM database provides information on diseases associated with the chromosomal location. Such associations include the results of linkage analysis mapped to this interval, and the correlation of translocations and other chromosomal aberrations in this area with the advent of polygenic diseases, such as cancer, in general and prostate cancer in particular.

E. Therapeutic applications of nucleic acid sequences

Nucleic acid sequences of the invention may also be used for therapeutic purposes. Turning first to the second aspect of the invention (i.e. inhibition of expression of variant), expression of variant product may be modulated through antisense technology, which controls gene expression through hybridization of complementary nucleic acid sequences, i.e. antisense DNA or RNA, to the control, 5' or regulatory regions of the gene encoding variant product. For example, the 5' coding portion of the nucleic acid sequence sequence which codes for the product of the present invention is used to design an antisense oligonucleotide of from about 10 to 40 base pairs in length. Oligonucleotides derived from the transcription start site, e.g. between positions -10 and +10 from the start site, are preferred. An antisense DNA oligonucleotide is designed to be

complementary to a region of the nucleic acid sequence involved in transcription (Lee *et al., Nucl. Acids, Res.*, **6**:3073, (1979); Cooney et al., Science **241**:456, (1988); and Dervan *et al., Science* **251**:1360, (1991)), thereby preventing transcription and the production of the variant products. An antisense RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into the variant products (Okano *J. Neurochem.* **56**:560, (1991)). The antisense constructs can be delivered to cells by procedures known in the art such that the antisense RNA or DNA may be expressed *in vivo*. The antisense may be antisense mRNA or DNA sequence capable of coding such antisense mRNA. The antisense mRNA or the DNA coding thereof can be complementary to the full sequence of nucleic acid sequences coding for the variant protein or to a fragment of such a sequence which is sufficient to inhibit production of a protein product.

Turning now to the first aspect of the invention, i.e. expression of variant, expression of variant product may be increased by providing coding sequences for coding for said product under the control of suitable control elements ending its expression in the desired host.

The nucleic acid sequences of the invention may be employed in combination with a suitable pharmaceutical carrier. Such compositions comprise a therapeutically effective amount of the compound, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The formulation should suit the mode of administration.

The products of the invention as well as any activators and deactivators compounds (see below) which are polypeptides, may also be employed in accordance with the present invention by expression of such polypeptides *in vivo*, which is often referred to as *"gene therapy."* Cells from a patient may be engineered with a nucleic acid sequence (DNA or RNA) encoding a polypeptide *ex vivo*, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, cells

may be engineered by procedures known in the art by use of a retroviral particle containing RNA encoding a polypeptide of the present invention.

Similarly, cells may be engineered in vivo for expression of a polypeptide *in vivo* by procedures known in the art. As known in the art, a producer cell for producing a retroviral particle containing RNA encoding the polypeptide of the present invention may be administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a product of the present invention by such method should be apparent to those skilled in the art from the teachings of the present invention. For example, the expression vehicle for engineering cells may be other than a retrovirus, for example, an adenovirus which may be used to engineer cells *in vivo* after combination with a suitable delivery vehicle.

Retroviruses from which the retroviral plasmid vectors mentioned above may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the *PE501*, *PA317*, *psi-2*, *psi-AM*, *PA12*, *T19-14X*, *VT-19-17-H2*, *psi-CRE*, *psi-CRIP*, *GP+E-86*, *GP+envAm12*, and *DAN* cell lines as described in Miller (*Human Gene Therapy*, Vol. **1**, pg. 5-14, (1990)). The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and $CaPO_4$ precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which include the nucleic acid sequence(s) encoding the polypeptides. Such retroviral vector particles then may be employed, to transduce eukaryotic cells,

either *in vitro* or *in vivo*. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

The genes introduced into cells may be placed under the control of inducible promoters, such as the radiation-inducible Egr-1 promoter, (Maceri, H.J., *et al., Cancer Res.*, **56**(19):4311 (1996)), to stimulate variant production or antisense inhibition in response to radiation, eg., radiation therapy for treating tumors.

Example IV. Variant product

The substantially purified variant product of the invention has been defined above as the product coded from the nucleic acid sequence of the invention. Preferably the amino acid sequence is an amino acid sequence having at least 90% identity to any one of the sequences identified as SEQ ID NO: 18 to SEQ ID NO: 34 provided that the amino acid sequence is not identical to that of the original sequence from which it has been varied. The protein or polypeptide may be in mature and/or modified form, also as defined above. Also contemplated are protein fragments having at least 10 contiguous amino acid residues, preferably at least 10-20 residues, derived from the variant product, as well as homologues as explained above.

The sequence variations are preferably those that are considered conserved substitutions, as defined above. Thus, for example, a protein with a sequence having at least 90% sequence identity with any of the products identified as SEQ ID NO: 18 to SEQ ID NO: 34, preferably by utilizing conserved substitutions as defined above is also part of the invention, and provided that it is not identical to the original peptide from which it has been varied. In a more specific embodiment, the protein has or contains any one of the sequence identified as SEQ ID NO: 18 to SEQ ID NO: 34. The variant product may be (i) one in which

one or more of the amino acid residues in a sequence listed above are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue), or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the variant product is fused with another compound, such as a compound to increase the half-life of the protein (for example, polyethylene glycol (PEG)), or a moiety which serves as targeting means to direct the protein to its target tissue or target cell population (such as an antibody), or (iv) one in which additional amino acids are fused to the variant product. Such fragments, variants and derivatives are deemed to be within the scope of those skilled in the art from the teachings herein.

A. Preparation of variant product

Recombinant methods for producing and isolating the variant product, and fragments of the protein are described above.

In addition to recombinant production, fragments and portions of variant product may be produced by direct peptide synthesis using solid-phase techniques (cf. Stewart *et al.*, (1969) Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco; Merrifield J., *J. Am. Chem. Soc.*, **85**:2149-2154, (1963)). In vitro peptide synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City, Calif.) in accordance with the instructions provided by the manufacturer. Fragments of variant product may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

B. Therapeutic uses and compositions utilizing the variant product

The variant product of the invention is generally useful in treating diseases and disorders which are characterized by a lower than normal level of variant expression, and or diseases which can be cured or ameliorated by raising the level of the variant product, even if the level is normal.

Variant products or fragments may be administered by any of a number of routes and methods designed to provide a consistent and predictable concentration of compound at the target organ or tissue. The product-containing compositions may be administered alone or in combination with other agents, such as stabilizing compounds, and/or in combination with other pharmaceutical agents such as drugs or hormones.

Variant product-containing compositions may be administered by a number of routes including, but not limited to oral, intravenous, intramuscular, transdermal, subcutaneous, topical, sublingual, or rectal means as well as by nasal application. variant product-containing compositions may also be administered via liposomes. Such administration routes and appropriate formulations are generally known to those of skill in the art.

The product can be given via intravenous or intraperitoneal injection. Similarly, the product may be injected to other localized regions of the body. The product may also be administered via nasal insufflation. Enteral administration is also possible. For such administration, the product should be formulated into an appropriate capsule or elixir for oral administration, or into a suppository for rectal administration.

The foregoing exemplary administration modes will likely require that the product be formulated into an appropriate carrier, including ointments, gels, suppositories. Appropriate formulations are well known to persons skilled in the art.

Dosage of the product will vary, depending upon the potency and therapeutic index of the particular polypeptide selected.

A therapeutic composition for use in the treatment method can include the product in a sterile injectable solution, the polypeptide in an oral delivery vehicle, the product in an aerosol suitable for nasal administration, or the product in a nebulized form, all prepared according to well known methods. Such compositions comprise a therapeutically effective amount of the compound, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not

limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The product of the invention may also be used to modulate endothelial differentiation and proliferation as well as to modulate apoptosis either *ex vivo* or *in vitro*, for example, in cell cultures.

Example V. Screening methods for activators and deactivators (inhibitors)

The present invention also includes an assay for identifying molecules, such as synthetic drugs, antibodies, peptides, or other molecules, which have a modulating effect on the activity of the variant product, e.g. activators or deactivators of the variant product of the present invention. Such an assay comprises the steps of providing an variant product encoded by the nucleic acid sequences of the present invention, contacting the variant protein with one or more candidate molecules to determine the candidate molecules modulating effect on the activity of the variant product, and selecting from the molecules a candidate's molecule capable of modulating variant product physiological activity.

The variant product, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening therapeutic compounds in any of a variety of drug screening techniques. The fragment employed in such a test may be free in solution, affixed to a solid support, borne on a cell membrane or located intracellularly. The formation of binding complexes, between variant product and the agent being tested, may be measured. Alternatively, the activator or deactivator may work by serving as agonist or antagonist, respectively, of the variant receptor, binding entity or target site, and their effect may be determined in connection with any of the above.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the variant product is described in detail by Geysen in PCT Application WO 84/03564, published on Sep. 13, 1984. In summary, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic

pins or some other surface. The peptide test compounds are reacted with the full variant product or with fragments of variant product and washed. Bound variant product is then detected by methods well known in the art. Substantially purified variant product can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

Antibodies to the variant product, as described in Example VI below, may also be used in screening assays according to methods well known in the art. For example, a *"sandwich"* assay may be performed, in which an anti-variant antibody is affixed to a solid surface such as a microtiter plate and variant product is added. Such an assay can be used to capture compounds which bind to the variant product. Alternatively, such an assay may be used to measure the ability of compounds to influence with the binding of variant product to the variant receptor, and then select those compounds which effect the binding.

Example VI. Anti-variant antibodies

A. Synthesis

In still another aspect of the invention, the purified variant product is used to produce anti-variant antibodies which have diagnostic and therapeutic uses related to the activity, distribution, and expression of the variant product.

Antibodies to the variant product may be generated by methods well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, humanized, single chain, Fab fragments and fragments produced by an Fab expression library. Antibodies, i.e., those which inhibit dimer formation, are especially preferred for therapeutic use.

A fragment of the variant product for antibody induction does not require biological activity but have to feature immunological activity; however, the protein fragment or oligopeptide must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least five

amino acids, preferably at least 10 amino acids of the sequences specified in any one of SEQ ID NO: 18 to SEQ ID NO: 34. Preferably they should mimic a portion of the amino acid sequence of the natural protein and may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of variant protein amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule. Procedures well known in the art can be used for the production of antibodies to variant product.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, etc may be immunized by injection with variant product or any portion, fragment or oligopeptide which retains immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include but are not limited to Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are potentially useful human adjuvants.

Monoclonal antibodies to variant protein may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Koehler and Milstein (*Nature* **256**:495-497, (1975)), the human B-cell hybridoma technique (Kosbor *et al., Immunol. Today* **4**:72, (1983); Cote *et al., Proc. Natl. Acad. Sci.* **80**:2026-2030, (1983)) and the EBV-hybridoma technique (Cole, *et al., Mol. Cell Biol.* 62:109-120, (1984)).

Techniques developed for the production of *"chimeric antibodies"*, the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can also be used (Morrison *et al., Proc. Natl. Acad. Sci.* **81**:6851-6855, (1984); Neuberger *et al., Nature* **312**:604-608, (1984); Takeda *et al., Nature* **314**:452-454, (1985)). Alternatively, techniques described for the production of single chain antibodies

(U.S. Pat. No. 4,946,778) can be adapted to produce single-chain antibodies specific for the variant protein.

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in Orlandi *et al.* (*Proc. Natl. Acad. Sci.* **86**:3833-3837, 1989)), and Winter G and Milstein C., (*Nature* **349**:293-299, (1991)).

Antibody fragments which contain specific binding sites for variant protein may also be generated. For example, such fragments include, but are not limited to, the $F(ab')_2$ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse W.D. *et al., Science* 256:1275-1281, (1989)).

B. Diagnostic applications of antibodies

A variety of protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the formation of complexes between the variant product and its specific antibody and the measurement of complex formation. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two noninterfering epitopes on a specific variant product is preferred, but a competitive binding assay may also be employed. These assays are described in Maddox D.E., *et al.,* (*J. Exp. Med.* **158**:1211, (1983)).

Antibodies which specifically bind variant product are useful for the diagnosis of conditions or diseases characterized by expression of the novel variant of the invention (where normally it is not expressed) by over or under expression of variant as well as for detection of diseases in which the proportion

between the amount of the variants of the invention and the original sequence from which it varied is altered. Alternatively, such antibodies may be used in assays to monitor patients being treated with variant product, its activators, or its deactivators. Diagnostic assays for variant protein include methods utilizing the antibody and a label to detect variant product in human body fluids or extracts of cells or tissues. The products and antibodies of the present invention may be used with or without modification. Frequently, the proteins and antibodies will be labeled by joining them, either covalently or noncovalently, with a reporter molecule. A wide variety of reporter molecules are known in the art.

A variety of protocols for measuring the variant product, using either polyclonal or monoclonal antibodies specific for the respective protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescent activated cell sorting (FACS). As noted above, a two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on variant product is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox, *et al. (supra)*. Such protocols provide a basis for diagnosing altered or abnormal levels of variant product expression. Normal or standard values for variant product expression are established by combining body fluids or cell extracts taken from normal subjects, preferably human, with antibody to variant product under conditions suitable for complex formation which are well known in the art. The amount of standard complex formation may be quantified by various methods, preferably by photometric methods. Then, standard values obtained from normal samples may be compared with values obtained from samples from subjects potentially affected by disease. Deviation between standard and subject values establishes the presence of disease state.

The antibody assays are useful to determine the level of variant product present in a body fluid sample, in order to determine whether it is being expressed at all, whether it is being overexpressed or underexpressed in the

tissue, or as an indication of how variant levels of variable products are responding to drug treatment.

By another aspect the invention concerns methods for determining the presence or level of various anti-variant antibodies in a biological sample obtained from patients, such as blood or serum sample using as an antigen the variant product. Determination of said antibodies may be indicative to a plurality of pathological conditions or diseases.

C. Therapeutic uses of antibodies

In addition to their diagnostic use the antibodies may have a therapeutical utility in blocking or decreasing the activity of the variant product in pathological conditions where beneficial effect can be achieved by such a decrease.

The antibody employed is preferably a humanized monoclonal antibody, or a human Mab produced by known globulin-gene library methods. The antibody is administered typically as a sterile solution by IV injection, although other parenteral routes may be suitable. Typically, the antibody is administered in an amount between about 1-15 mg/kg body weight of the subject. Treatment is continued, e.g., with dosing every 1-7 days, until a therapeutic improvement is seen.

Although the invention has been described with reference to specific methods and embodiments, it is appreciated that various modifications and changes may be made without departing from the invention.

CLAIMS:

1. An isolated nucleic acid sequence, of an alternative splicing variant, selected from the group consisting of:

(i) the nucleic acid sequence depicted in any one of SEQ ID NO: 1 to SEQ ID NO: 17;

(ii) nucleic acid sequences having at least 90% identity with the sequence of (i) with the proviso that each sequence is different than the original nucleic acid sequence from which the sequences of (i) have been varied by alternative splicing; and

(iii) fragments of (i) or (ii) of at least 20 b.p., provided that said fragment contains a sequence which is not present, as a continuous stretch of nucleotides, in the original nucleic acid sequence from which the sequences of (i) have been varied by alternative splicing.

2. An isolated nucleic acid sequence complementary to the nucleic acid sequence of Claim 1.

3. An amino acid sequence selected from the group consisting of:

(i) an amino acid sequence coded by the isolated nucleic acid sequence of alternative splice variants of Claim 1;

(ii) homologues of the amino acid sequences of (i) in which one or more amino acids has been added, deleted, replaced or chemically modified in the region or adjacent to the region where the amino acid sequences differs from the original amino acid sequence, coded by the original nucleic acid sequence from which the variant has been varied.

4. An amino acid sequence according to Claim 3, as depicted in any one of SEQ ID NO: 18 to SEQ ID NO: 34.

5. An isolated nucleic acid sequence coding for any one of the amino acid sequences of Claim 3 or 4.

6. A purified antibody which binds specifically to any of the amino acid sequence of Claim 3 or 4.

7. An expression vector comprising any one of the nucleic acid sequences of Claim 1 or 5 and control elements for the expression of the nucleic acid sequence in a suitable host.

8. An expression vector comprising any one of the nucleic acid sequences of Claim 2, and control elements for the expression of the nucleic acid sequences in a suitable host.

9. A host cell transfected by the expression vector of Claim 7 or 8.

10. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and as an active ingredient an agent selected from the group consisting of:

(i) the expression vector of Claim 7; and

(ii) any one of the amino acid sequences of Claim 3 or 4.

11. A pharmaceutical composition according to Claim 10, for treatment of diseases which can be ameliorated or cured by raising the level of any one of the amino acid sequences depicted in SEQ ID NO: 18 to SEQ ID NO: 34.

12. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and as an active ingredient an agent selected from the group consisting of:

(i) any one of the nucleic acid sequences of Claim 2;

(ii) the expression vector of Claim 8; and

(iii) the purified antibody of Claim 6.

13. A pharmaceutical composition according to Claim 12, for treatment of diseases which can be ameliorated or cured by decreasing the level of any one of the amino acid sequences depicted in SEQ ID NO: 18 to SEQ ID NO: 34.

14. A method for detecting an variant nucleic acid sequence in a biological sample, comprising the steps of:

(a) hybridizing to nucleic acid material of said biological sample any one of the nucleic acid sequences of Claim 1 or 2; and

(b) detecting said hybridization complex;

wherein the presence of said hybridization complex correlates with the presence of an variant nucleic acid sequence in the said biological sample.

15. A method for determining the level of variant nucleic acid sequences in a biological sample comprising the steps of:

(a) hybridizing to nucleic acid material of said biological sample any one of the nucleic acid sequences of Claim 1 or 2; and

(b) determining the amount of hybridization complexes and normalizing said amount to provide the level of the variant nucleic acid sequences in the sample.

16. A method for determining the ratio between the level of variant of the nucleic acid sequence in a first biological sample and the level of the original sequence from which the variant has been varied by alternative splicing in a second biological sample comprising:

(a) determining the level of the variant nucleic acid sequence in the first biological sample according to the method of Claim 15;

(b) determining the level of the original sequence in the second biological sample; and

(c) comprising the levels obtained in (a) and (b) to give said ratio.

17. A method according to Claim 16, wherein said first and said second biological samples are the same sample.

18. A method according to any of Claims 14 to 17, wherein the nucleic acid material of said biological sample are mRNA transcripts.

19. A method according to Claim 18, where the nucleic acid sequence is present in a nucleic acid chip.

20. A method for identifying candidate compounds capable of binding to the variant product and modulating its activity the method comprising:

(i) providing any one of the amino acid sequences as defined in Claim 3 or 4;

(ii) contacting a candidate compound with said amino acid sequence;

(iii) determining the effect of said candidate compound on the biological activity of said protein or polypeptide and selecting those compounds which show a significant effect on said biological activity.

21. A method according to Claim 20, wherein the compound is an activator and the measured effect is increase in the biological activity.

22. A method according to Claim 20, wherein the compound is an deactivator and the effect is decrease in the biological activity.

23. An activator of any one of the amino acid sequences of Claim 3 or 4.

24. An deactivator of any one of the amino acid sequences of Claims 3 or 4.

25. A method for detecting any one of the amino acid sequences of Claim 3 or 4 in a biological sample, comprising the steps of:

(a) contacting with said biological sample the antibody of Claim 8, thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

wherein the presence of said antibody-antigen complex correlates with the presence of the desired amino acid in said biological sample.

26. A method for detecting the level of the amino acid sequence of any one of Claim 3 or 4 in a biological sample, comprising the steps of:

(a) contacting with said biological sample the antibody of Claim 8, thereby forming an antibody-antigen complex; and

(b) detecting the amount of said antibody-antigen complex and normalizing said amount to provide the level of said amino acid sequence in the sample.

27. A method for determining the ratio between the level of any one of the amino acid sequences of Claims 3 or 4 present in a first biological sample and the level of the original amino acid sequences from which they were varied by alternative splicing, present in a second biological sample, the method comprising:

(a) determining the level of the amino acid sequences of Claims 3 or 4 into a first sample by the method of Claim 26;

(b) determining the level of the original amino acid sequence in the second sample; and

(c) comparing the level obtained in (a) and (b) to give said ratio.

28. A method according to Claim 27, wherein said first and said second biological samples are the same sample.

29. A method for detecting any one of the antibodies of Claim 6 in a biological sample comprising the steps of:

(a) contacting said biological sample with any one of the amino acid sequences of Claim 3 or 4 thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

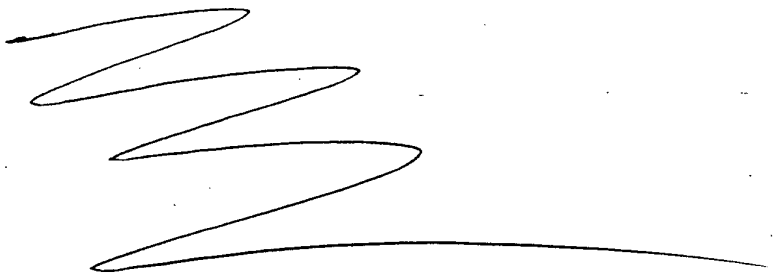
wherein the presence of said antibody-antigen complex correlates with the presence of the antibody in said biological sample.

30. A method for detecting the level of any one of the antibodies of Claim 6 in a biological sample comprising the steps of:

(a) contacting said biological sample with any one of the amino acid sequences of Claim 3;

(b) detecting the amount of said antibody-antigen complex and normalizing said amount to provide the levels of said antibody in the sample.

For the Applicants,
REINHOLD COHN AND PARTNERS
By:



SEQUENCE LISTING

<110> COMPUGEN LTD

<120> ALTERNATICE SPLICE VARIANTS

<130> COMPUGEN6

<140>
<141>

<160> 34

<170> PatentIn Ver. 2.1

<210> 1
<211> 1530
<212> DNA
<213> HUMAN

<400> 1
caggccttga ggttttggca gctctggagg atgagagaga acatggccag gggcccttgc 60
aacgcgccga gatgggtgtc cctgatggtg ctcgtcgcca taggcaccgc cgtgacagcg 120
gccgtcaacc ctggcgtcgt ggtcaggatc tcccagaagg gcctggacta cgccagccag 180
caggggacgg ccgctctgca gaaggagctg aagaggatca agattcctga ctactcagac 240
agctttaaga tcaagcatct tgggaagggg cattatagct tctacagcat ggacatccgt 300
gaattccagc ttcccagttc ccagataagc atggtgccca atgtgggcct taagttctcc 360
atcagcaacg ccaatatcaa gatcagcggg aaatggaagg cacaaaagag attcttaaaa 420
atgagcggca attttgacct gagcatagaa ggcatgtcca tttcggctga tctgaagctg 480
ggcagtaacc ccacgtcagg caagcccacc atcacctgct ccagctgcag cagccacatc 540
aacagtgtcc acgtgcacat ctcaaagagc aaagtcgggt ggctgatcca actcttccac 600
aaaaaaattg agtctgcgct tcgaaacaag atgaacagcc aggtctgcga gaaagtgacc 660
aattctgtat cctccaagct gcaaccttat ttccagactc tgccagtaat gaccaaaata 720
gattctgtgg ctggaatcaa ctatggtctg gtggcacctc cagcaaccac ggctgagacc 780
ctggatgtac agatgaaggg ggagttttac agtgagaacc accacaatcc acctcccttt 840
gctccaccag tgatggagtt tcccgctgcc catgaccgca tggtatacct gggcctctca 900
gactacttct tcaacacagc cgggcttgta taccaagagg ctggggtctt gaagatgacc 960
cttagagatg acatgattcc aaaggagtcc aaatttcgac tgacaaccaa gttctttgga 1020
accttcctac ctgaggtggc caagaagttt cccaacatga agatacagat ccatgtctca 1080
gcctccaccc cgccacacct gtctgtgcag cccaccggcc ttaccttcta ccctgccgtg 1140
gatgtccagg cctttgccgt cctccccaac tcctccctgg cttccctctt cctgattggc 1200
atggtaagca gttcctgggt tggacagatg aggagcccca gacagtccca acagcactgt 1260
ctttggagtc aggagaccat gtgaatcctg tctggattca aacctggact gtgtcactcc 1320
ggagcctgag gcttgagtca ctgtactcaa tggtgccgac tcctggaggt attcattcac 1380
ccagccattc actagtgcgt ttgtttactt attcattcaa ttattcattc agtcaatttc 1440
tcattcattc anttattcat tccatgttgg cttgaaatat gtgtactgtn ccaattnatc 1500
catttatatc tttagtcatt caattatgca 1530

<210> 2
<211> 2251
<212> DNA
<213> HUMAN

<400> 2
ttttttcagg ggtatgaggc tttattcagg aagagctcaa gatgggaggg gagcccctgg 60
ggtcccactc ttgtggcctc ttgataccaa ggctcatggg acgctaaaac ctaatgactg 120
caggcatctg agcttccttg gagtcatctg tgggggcac aggggcagcc tcagagtcac 180
ccgtgggggg cacaggggcg gccccggagt cacccgtggg cggcacgggg gcggtctcgg 240
agtcacccgt agcggggatg gggcagcctc agtgtcaccc aaggtgggca cgggcctcag 300
agtcacccgt gggcggcacg gggggggccc cggagtcacc cgtgggcggc acgggggggg 360

ccccggagtc acccgtgggc ggcacggggg gggccccgga gtcacccgtg ggcggcacgg 420
gggcggtctc ggagtcaccc gtgggggggca cgggagtggc ctcggagtcc cctgtggggg 480
gcacaggggt ggcctcctgg tcggtcactg tgggcagcgc cagataggtg aggctccagt 540
agcgcaggaa gttggttctc aggctccgct tcatggagct gctgcccatc ttcttggtga 600
tctccaggta gccgctgttt tccgtagtgt agggttccca gtgtgtgggc acagccgagt 660
cgcccatgtt ggggtcccct gttttggcaa agttggtcca gtaggcgatc atggccttag 720
agactgtcct gtcttggggc cggtagcccg tgggggtggc gaagggcttc ccgaaaacgt 780
actgaatgtc atctgcatgg tcggccccca cccatttggg gtagacgggc atccgagagg 840
gatgggaaaa caggtaggcg taggtcttgg cactcttggc attggctctg tgctgggcta 900
gggcaatctc ggtgggcacc aggaagagga catcggtctc aaagtccacc acagtcttct 960
tcttattctc ctgggatggg tcctgggccc aggactcggt gtagacatca aaggtcgtct 1020
tggcgcctct gagccccttg gtgattgtga actcactgac cagcttgtag aagtcctcct 1080
ccgtgacttt cttgttgccc ttgttgatgg caggcatgtc gatgctggcg aagatgtggc 1140
cgtccatgtt gttggtgcct gctatatagt cgatgtcggc ggcgttggcg tacaggttga 1200
tcgggtcagc ggggatgaag tctccatcaa tgacagggac gaagcccaca tagtgcagca 1260
tggggtactc caggcctgcc agcggcacct tataggccag cgtcagggct cggggatcag 1320
taaccttcag acactgggcc atcctggcgg catcacccac agggcaaccc accttctcag 1380
ccacctttt ggcccagaag agtgggtttt tctggatgac ccagggactc agggccacgc 1440
cgctctggct gatggctcgc cggatgaggc ccttgttgta gggggagagg gtctgcagag 1500
agacgctggc acctccagca gactccccga agagcgtgat gttgttgggg tccccccga 1560
aggccgcgat attcctcttc acccaagcaa tggccatgtg ctgatcccga agaccatagt 1620
tacctggcag attggcgtcc ccagtgctga ggaacccaag ggggccgaca cggtagttga 1680
aggtgaccac gatgacgttt ccgcgtgtgg cgatctcctc gccgtcatac aggtagttgt 1740
tgaggaagtt ggccccatgg ccggacccca tgaggaaggc gcctccatag atccagatca 1800
taacgggcag gtcccgggag acttgcttcc tgccctgggg cacccaaatg ttgaggtaca 1860
ggcagtcttc atccccgtag gtgctgtcct gggtgatggt ggcctgcagg catctcttct 1920
tgaagttctt ggccttcagg gtcccttgcc agccaggatg tggctgagga ttttccaggg 1980
ccttggtggg agctgcgaag gggatgccct tgaagatgtc cacagagtca cccaggaggc 2040
cgagcttctt attgacgcct tccacgaacc caccttctgt gtacacggcg cccagcttcg 2100
cggcactcgc cactgcccag cagcaggtga ggcccaacac aaccagttgc aggcgcccca 2160
tggtgagcat cagcctctgg atgaataaac tctttgcatt atcttgagac tgctgttcct 2220
tccatcttga cacttaactg atacagaagg t 2251

<210> 3
<211> 2078
<212> DNA
<213> HUMAN

<400> 3
ctgctgctcc tggggttcct gctggtgagc ttggagtcaa cactttcgat tccaccttgg 60
gaagccccca aggagcataa gtacaaagct gaagagcaca cagtcgttct cactgtcacc 120
ggggagccct gccacttccc cttccagtac caccggcagc tgtaccacaa atgtacccac 180
aagggccggc caggccctca gccctggtgt gctaccaccc ccaactttga tcaggaccag 240
cgatggggat actgtttgga gcccaagaaa gtgaaagacc actgcagcaa acacagcccc 300
tgccagaaag gagggacctg tgtgaacatg ccaagcggcc cccactgtct ctgtccacaa 360
cacctcactg gaaaccactg ccagaaagag aagtgctttg agcctcagct tctccggttt 420
ttccacaaga atgagatatg gtatagaact gagcaagcag ctgtggccag atgccagtgc 480
aagggtcctg atgcccactg ccagcggctg gccagccagg cctgccgcac caacccgtgc 540
ctccatgggg gtcgctgcct agaggtggag ggccaccgcc tgtgccactg cccggtgggc 600
tacaccggac ccttctgcga cgtggacacc aaggcaagct gctatgatgg ccgcgggctc 660
agctaccgcg gcctggccag gaccacgctc tcgggtgcgc cctgtcagcc gtgggcctcg 720
gaggccacct accggaacgt gactgccgag caagcgcgga actggggact gggcggccac 780
gccttctgcc ggaacccgga caacgacatc cgcccgtggt gcttcgtgct gaaccgcgac 840
cggctgagct gggagtactg cgacctggca cagtgccaga ccccaaccca ggcggcgcct 900
ccgaccccgg tgtcccctag gcttcatgtc ccactcatgc ccgcgcagcc ggcaccgccg 960
aagcctcagc ccacgacccg gaccccgcct cagtcccaga ccccgggagc cttgccggcg 1020
aagcgggagc agccgccttc cctgaccagg aacggcccac tgagctgcgg gcagcggctc 1080
cgcaagagtc tgtcttcgat gacccgcgtc gttggcgggc tggtggcgct acgcggggcg 1140
cacccctaca tcgccgcgct gtactggggc cacagtttct gcgccggcag cctcatcgcc 1200
ccctgctggg tgctgacggc cgctcactgc ctgcaggacc ggcccgcacc cgaggatctg 1260

acggtggtgc tcggccagga acgccgtaac cacagctgtg agccgtgcca gacgttggcc 1320
gtgcgctcct accgcttgca cgaggccttc tcgcccgtca gctaccagca cgacctggct 1380
ctgttgcgcc ttcaggagga tgcggacggc agctgcgcgc tcctgtcgcc ttacgttcag 1440
ccggtgtgcc tgccaagcgg cgccgcgcga ccctccgaga ccacgctctg ccaggtggcc 1500
ggctggggcc accagttcga gggggcggag gaatatgcca gcttcctgca ggaggcgcag 1560
gtaccgttcc tctccctgga gcgctgctca gccccggacg tgcacggatc ctccatcctc 1620
cccggcatgc tctgcgcagg gttcctcgag ggcggcaccg atgcgtgcag gtgagctctt 1680
agccggttgg cgcccttccc cgaggccgtc agcawaatct caggtccaca gcgctgactg 1740
cgtgtttccg acccagggtg attccggagg cccgctggtg tgtgaggacc aagctgcaga 1800
gcgccggctc accctgcaag gcatcatcag ctggggatcg ggctgtggtg accgcaacaa 1860
gccaggcgtc tacaccgatg tggcctacta cctggcctgg atccgggagc acaccgtttc 1920
ctgattgctc agggactcat ctttccctcc ttggtgattc cgcagtgaga gagtggctgg 1980
ggcatggaag gcaagattgt gtcccattcc cccagtgcgg ccagctccgc gccaggatgg 2040
cgcaggaact caataaagtg ctttgaaaat gctgagaa 2078

<210> 4
<211> 720
<212> DNA
<213> HUMAN

<400> 4
tcgaccgccc cttggcttct gcactgatgg tgggtggatg agtaatgcat ccaggaagcc 60
tggaggcctg tggtttccgc acccgctgcc acccccgccc ctagcgtgga catttatcct 120
ctagcgctca ggccctgcgc gcatcgccgc agatccagcg cccagagaga caccagagaa 180
cccaccatgg ccccctttga gcccctggct tctggcatcc tgttgttgct gtggctgata 240
gcccccagca gggcctgcac ctgtgtccca ccccacccac agacggcctt ctgcaattcc 300
gacctcgtca tcagggccaa gttcgtgggg acaccagaag tcaaccagac caccttatac 360
cagcgttatg agatcaagat gaccaagatg tataaagggt tccaagcctt aggggatgcc 420
gctgacatcc ggttcgtcta cacccccgcc atggagagtg tctgcggata cttccacagg 480
tcccacaacc gcagcgagga gtttctcatt gctggaaaac tgcaggatgg actcttgcac 540
atcactacct gcagttttgt ggctccctgg aacagcctga gcttagctca gcgccggggc 600
ttcaccaaga cctacactgt tggctgtgag gaatgcacag tgtttccctg ttcccactcc 660
catctttctt ccggacaatg aaataaagag ttaccaccca gcagaaaaaa caaacaagtc 720

<210> 5
<211> 556
<212> DNA
<213> HUMAN

<400> 5
tcgaccgccc cttggcttct gcactgatgg tgggtggatg agtaatgcat ccaggaagcc 60
tggaggcctg tggtttccgc acccgctgcc acccccgccc ctagcgtgga catttatcct 120
ctagcgctca ggccctgcgc gcatcgccgc agatccagcg cccagagaga caccagagaa 180
cccaccatgg ccccctttga gcccctggct tctggcatcc tgttgttgct gtggctgata 240
gcccccagca gggcctgcac ctgtgtccca ccccacccac agacggcctt ctgcaattcc 300
gacctcgtca tcagggccaa gttcgtgggg acaccagaag tcaaccagac caccttatac 360
cagcgttatg agatcaagat gaccaagatg tataaagggt tccaagcctt aggggatgcc 420
gctgacatcc ggttcgtcta cacccccgcc atggagagtg tctgcggata cttccacagg 480
tcccacaacc gcagcgagga gtttctcatt gctggaaaac tgcaggtagt gatgtgcaag 540
agtccatccg tcgtac 556

<210> 6
<211> 934
<212> DNA
<213> HUMAN

<400> 6
tcgaccgccc cttggcttct gcactgatgg tgggtggatg agtaatgcat ccaggaagcc 60

tggaggcctg tggtttccgc acccgctgcc acccccgccc ctagcgtgga catttatcct 120
ctagcgctca ggccctgcgc gcatcgccgc agatccagcg cccagagaga caccagagaa 180
cccaccatgg ccccctttga gcccctggct tctggcatcc tgttgttgct gtggctgata 240
gcccccagca gggcctgcac ctgtgtccca ccccacccac agacggcctt ctgcaattcc 300
gacctcgtca tcagggccaa gttcgtgggg acaccagaag tcaaccagac caccttatac 360
cagcgttatg agatcaagat gaccaagatg tataaagggt tccaagcctt aggggatgcc 420
gctgacatcc ggttcgtcta cacccccgcc atggagagtg tctgcggata cttccacagg 480
tcccacaacc gcagcgagga gtttctcatt tgctgggaaa attgcaggat gggatttttg 540
cacattcact aacctgcagt ttttgttggg ttccttggga acagcctgag cttagctcag 600
cgccggggct tcaccaagac ctacactgtt ggctgtgagg aatgcacagt gtttccctgt 660
ttatccatcc cctgcaaact gcagagtggc actcattgct tgtggacgga ccagctcctc 720
caaggctctg aaaagggctt ccagtcccgt caccttgcct gcctgcctcg ggagccaggg 780
ctgtgcacct ggcagtccct gcggtcccag atagcctgaa tcctgcccgg agtggaagct 840
gaagcctgca cagtgtccac cctgttccca ctcccatctt tcttccggac aatgaaataa 900
agagttacca cccagcagaa aaaacaaaca agtc 934

<210> 7
<211> 783
<212> DNA
<213> HUMAN

<400> 7
tcgaccgccc cttggcttct gcactgatgg tgggtggatg agtaatgcat ccaggaagcc 60
tggaggcctg tggtttccgc acccgctgcc acccccgccc ctagcgtgga catttatcct 120
ctagcgctca ggccctgcgc gcatcgccgc agatccagcg cccagagaga caccagagaa 180
cccaccatgg ccccctttga gcccctggct tctggcatcc tgttgttgct gtggctgata 240
gcccccagca gggcctgcac ctgtgtccca ccccacccac agacggcctt ctgcaattcc 300
gacctcgtca tcagggccaa gttcgtgggg acaccagaag tcaaccagac caccttatac 360
cagcgttatg agatcaagat gaccaagatg tataaagggt tccaagcctt aggggatgcc 420
gctgacatcc ggttcgtcta cacccccgcc atggagagtg tctgcggata cttccacagg 480
tcccacaacc gcagcgagga gtttctcatt tatccatccc ctgcaaactg cagagtggca 540
ctcattgctt gtggacggac cagctcctcc aaggctctga aaagggcttc cagtcccgtc 600
accttgcctg cctgcctcgg gagccagggc tgtgcacctg gcagtccctg cggtcccaga 660
tagcctgaat cctgcccgga gtggaagctg aagcctgcac agtgtccacc ctgttcccac 720
tcccatcttt cttccggaca atgaaataaa gagttaccac ccagcagaaa aaacaaacaa 780
gtc 783

<210> 8
<211> 895
<212> DNA
<213> HUMAN

<400> 8
tcgaccgccc cttggcttct gcactgatgg tgggtggatg agtaatgcat ccaggaagcc 60
tggaggcctg tggtttccgc acccgctgcc acccccgccc ctagcgtgga catttatcct 120
ctagcgctca ggccctgcgc gcatcgccgc agatccagcg cccagagaga caccagagaa 180
cccaccatgg ccccctttga gcccctggct tctggcatcc tgttgttgct gtggctgata 240
gcccccagca gggcctgcac ctgtgtccca ccccacccac agacggcctt ctgcaattcc 300
gacctcgtca tcagggccaa gttcgtgggg acaccagaag tcaaccagac caccttatac 360
cagcgttatg agatcaagat gaccaagatg tataaagggt tccaagcctt aggggatgcc 420
gctgacatcc ggttcgtcta cacccccgcc atggagagtg tctgcggata cttccacagg 480
gctggaaaac tgcaggatgg actcttgcac atcactacct gcagttttgt ggctccctgg 540
aacagcctga gcttagctca gcgccggggc ttcaccaaga cctacactgt tggctgtgag 600
gaatgcacag tgtttccctg tttatccatc ccctgcaaac tgcagagtgg cactcattgc 660
ttgtggacgg accagctcct ccaaggctct gaaaagggct tccagtcccg tcaccttgcc 720
tgcctgcctc gggagccagg gctgtgcacc tggcagtccc tgcggtccca gatagcctga 780
atcctgcccg gagtggaagc tgaagcctgc acagtgtcca ccctgttccc actcccatct 840
ttcttccgga caatgaaata aagagttacc acccagcaga aaaaacaaac aagtc 895

<210> 9
<211> 2565
<212> DNA
<213> HUMAN

<400> 9
ggcacgagca gccccggagc ccgggccagg gtccacctgt ccccgcagcg ccggctcgcg 60
ccctcctgcc gcagccaccg agccgccgtc tagcgccccg acctcgccac catgagagcc 120
ctgctggcgc gcctgcttct ctgcgtcctg gtcgtgagcg actccaaagg cagcaatgaa 180
cttcatcaag ttccatcgaa ctgtgactgt ctaaatggag gaacatgtgt gtccaacaag 240
tacttctcca acattcactg gtgcaactgc ccaaagaaat tcggagggca gcactgtgaa 300
atagataagt caaaaacctg ctatgagggg aatggtcact tttaccgagg aaaggccagc 360
actgacacca tgggccggcc ctgcctgccc tggaactctg ccactgtcct tcagcaaacg 420
taccatgccc acagatctga tgctcttcag ctgggcctgg ggaaacataa ttactgcagg 480
gaggtggggg cacaaggacc aaaagccctc cctacagttc ccagaaacct tgttaccatc 540
cccttctccc agagggctgg ccatagcaca agagaagtgc agcctctggt tgagtcttcc 600
ctgaggggag gaggcaggga aggccctctg ggttggaatg acatccccta tctttctgtg 660
ttgccaggaa cccagacaac cggaggcgac cctggtgcta tgtgcaggtg ggcctaaagc 720
cgcttgtcca agagtgcatg gtgcatgact gcgcagatgg aaaaaagccc tcctctcctc 780
cagaagaatt aaaatttcag tgtggccaaa agactctgag gccccgcttt aagattattg 840
ggggagaatt caccaccatc gagaaccagc cctggtttgc ggccatctac aggaggcacc 900
gggggggctc tgtcacctac gtgtgtggag gcagcctcat cagcccttgc tgggtgatca 960
gcgccacaca ctgcttcatt gattacccaa agaaggagga ctacatcgtc tacctgggtc 1020
gctcaaggct taactccaac acgcaagggg agatgaagtt tgaggtggaa aacctcatcc 1080
tacacaagga ctacagcgct gacacgcttg ctcaccacaa cgacattgcc ttgctgaaga 1140
tccgttccaa ggagggcagg tgtgcgcagc catcccggac tatacagacc atctgcctgc 1200
cctcgatgta taacgatccc cagtttggca caagctgtga gatcactggc tttggaaaag 1260
agaattctac cgactatctc tatccggagc agctgaaaat gactgttgtg aagctgattt 1320
cccaccggga gtgtcagcag ccccactact acggctctga agtcaccacc aaaatgctgt 1380
gtgctgctga cccacagtgg aaaacagatt cctgccaggg agactcaggg ggacccctcg 1440
tctgttccct ccaaggccgc atgactttga ctggaattgt gagctggggc cgtggatgtg 1500
ccctgaagga caagccaggc gtctacacga gagtctcaca cttcttaccc tggatccgca 1560
gtcacaccaa ggaagagaat ggcctggccc tctgagggtc cccagggagg aaacgggcac 1620
cacccgcttt cttgctggtt gtcatttttg cagtagagtc atctccatca gctgtaagaa 1680
gagactggga agataggctc tgcacagatg gatttgcctg tgccaccacc agggygaacg 1740
acaatagctt taccctcagg cataggcctg ggtgctggct gcccagaccc ctctggccag 1800
gatggagggg tggtcctgac tcaacatgtt actgaccagc aacttgtctt tttctggact 1860
gaagcctgca ggagttaaaa agggcagggc atctcctgtg catgggtgaa gggagagcca 1920
gctcccccga cggtgggcat ttgtgaggcc catggttgag aaatgaataa tttcccaatt 1980
aggaagtgta acagctgagg tctcttgagg gagcttagcc aatgtgggag cagcggtttg 2040
gggagcagag acactaacga cttcagggca gggctctgat attccatgaa tgtatcagga 2100
aatatatatg tgtgtgtatg tttgcacact tgtgtgtggg ctgtgagtgt aagtgtgagt 2160
aagagctggt gtctgattgt taagtctaaa tatttcctta aactgtgtgg actgtgatgc 2220
cacacagagt ggtctttctg gagaggttat aggtcactcc tggggcctct tgggtccccc 2280
acgtgacagt gcctgggaat gtattattct gcagcatgac ctgtgaccag cactgtctca 2340
gtttcacttt cacatagatg tccctttctt ggccagttat cccttccttt tagcctagtt 2400
catccaatcc tcactgggtg gggtgaggac cactccttac actgaatatt tatatttcac 2460
tatttttatt tatatttttg taattttaaa taaaagtgat caataaaatg tgatttttct 2520
gatgamaaaa aaaaaagggc ggccgctcta gagtatccct cgagg 2565

<210> 10
<211> 2438
<212> DNA
<213> HUMAN

<400> 10
ggggcacagc tgtgtttggc tgcagggcca agagcgctgt caagaagacc cacacgcccc 60
cctccagcag ctgaattcct gcagctcagc agccgccgcc agagcaggac gaaccgccaa 120
tcgcaaggca cctctgagaa cttcaggatg cagatgtctc cagccctcac ctgcctagtc 180

ctgggcctgg cccttgtctt tggtgaaggg tctgctgtgc accatccccc atcctacgtg 240
gcccacctgg cctcagactt cggggtgagg gtgtttcagc aggtggcgca ggcctccaag 300
gaccgcaacg tggttttctc accctatggg gtggcctcgg tgttggccat gctccagctg 360
acaacaggag gagaaaccca gcagcagatt caagcagcta tgggattcaa gattgatgac 420
aagggcatgg cccccgccct ccggcatctg tacaaggagc tcatggggcc atggaacaag 480
gatgagatca gcaccacaga cgcgatcttc gtccagcggg atctgaagct ggtccagggc 540
ttcatgcccc acttcttcag gctgttccgg agcacggtca agcaagtgga cttttcagag 600
gtggagagag ccagattcat catcaatgac tgggtgaaga cacacacaaa aggtatgatc 660
agcaacttgc ttgggaaagg agccgtggac cagctgacac ggctggtgct ggtgaatgcc 720
ctctacttca acggccagtg gaagactccc ttccccgact ccagcaccca ccgccgcctc 780
ttccacaaat cagacggcag cactgtctct gtgcccatga tggctcagac caacaagttc 840
aactatactg agttcaccac gcccgatggc cattactacg acatcctgga actgccctac 900
cacggggaca ccctcagcat gttcattgct gcagacctgg ttcccactga ggccctttgc 960
aggatggaac tacggggctt acaggagctt ttgtgtgcct ggtagaaact atttctgttc 1020
cagtcacatt gccatcactc ttgtactgcc tgccaccgcg gaggaggctg gtgacaggcc 1080
aaaggccagt ggaagaaaca ccctttcatc tcagagtcca ctgtggcact ggccacccct 1140
ccccagtaca ggggtgctgc aggtggcaga gtgaatgtcc cccatcatgt ggcccaactc 1200
tcctggcctg gccatctccc tccccagaaa cagtgtgcat gggttatttt ggagtgtagg 1260
tgacttgttt actcattgaa gcagatttct gcttcctttt atttttatag gaatagagga 1320
agaaaggtca gatgcgtgcc cagctcttca ccccccaatc tcttggtggg gaggggtgta 1380
cctaaatatt tatcatatcc ttgcccttga gtgcttgtta gagagaaaga gaactactaa 1440
ggaaaataat attatttaaa ctcgctccta gtgtttcttt gtggtctgtg tcaccgtatc 1500
tcaggaagtc cagccacttg actggcacac acccctccgg acatccagcg tgacggagcc 1560
cacactgcca ccttgtggcc gcctgagacc ctcgcgcccc ccgcgccccc cgcgcccctc 1620
tttttccccc tgatggaaat tgaccataca atttcatcct ccttcagggg atcaaaagga 1680
cggagtgggg ggacagagac tcagatgagg acagagtggt ttccaatgtg ttcaatagat 1740
ttaggagcag aaatgcaagg ggctgcatga cctaccagga cagaactttc cccaattaca 1800
gggtgactca cagccgcatt ggtgactcac ttcaatgtgt catttccggc tgctgtgtgt 1860
gagcagtgga cacgtgaggg gggggtgggt gagagagaca ggcagctcgg attcaactac 1920
cttagataat atttctgaaa acctaccagc cagagggtag ggcacaaaga tggatgtaat 1980
gcactttggg aggccaaggc gggaggattg cttgagccca ggagttcaag accagcctgg 2040
gcaacatacc aagacccccg tctctttaaa aatatatata ttttaaatat acttaaatat 2100
atatttctaa tatctttaaa tatatatata tattttaaag accaatttat gggagaattg 2160
cacacagatg tgaaatgaat gtaatctaat agaagcctaa tcagcccacc atgttctcca 2220
ctgaaaaatc ctctttcttt ggggtttttc tttctttctt ttttgatttt gcactggacg 2280
gtgacgtcag ccatgtacag gatccacagg ggtggtgtca aatgctattg aaattgtgtt 2340
gaattgtatg ctttttcact tttgataaat aaacatgtaa aaatgtttca aaaaaataat 2400
aaaataaata aatacgaaga atatgtcagg acaaaaaa 2438

<210> 11
<211> 2522
<212> DNA
<213> HUMAN

<400> 11
ggggcacagc tgtgtttggc tgcagggcca agagcgctgt caagaagacc cacacgcccc 60
cctccagcag ctgaattcct gcagctcagc agccgccgcc agagcaggac gaaccgccaa 120
tcgcaaggca cctctgagaa cttcaggatg cagatgtctc cagccctcac ctgcctagtc 180
ctgggcctgg cccttgtctt tggtgaaggg tctgctgtgc accatccccc atcctacgtg 240
gcccacctgg cctcagactt cggggtgagg gtgtttcagc aggtggcgca ggcctccaag 300
gaccgcaacg tggttttctc accctatggg gtggcctcgg tgttggccat gctccagctg 360
acaacaggag gagaaaccca gcagcagatt caagcagcta tgggattcaa gattgatgac 420
aagggcatgg cccccgccct ccggcatctg tacaaggagc tcatggggcc atggaacaag 480
gatgagatca gcaccacaga cgcgatcttc gtccagcggg atctgaagct ggtccagggc 540
ttcatgcccc acttcttcag gctgttccgg agcacggtca agcaagtgga cttttcagag 600
gtggagagag ccagattcat catcaatgac tgggtgaaga cacacacaaa aggtatgatc 660
agcaacttgc ttgggaaagg agccgtggac cagctgacac ggctggtgct ggtgaatgcc 720
ctctacttca acggccagtg gaagactccc ttccccgact ccagcaccca ccgccgcctc 780
ttccacaaat cagacggcag cactgtctct gtgcccatga tggctcagac caacaagttc 840
aactatactg agttcaccac gcccgatggc cattactacg acatcctgga actgccctac 900

cacggggaca ccctcagcat gttcattgct gccccttatg aaaaagaggt gcctctctct 960
gccctcacca acattctgag tgcccagctc atcagccact ggaaaggcaa catgaccagg 1020
ctgccccgcc tcctggttct gcccaagttc tccctggaga ctgaagtcga cctcaggaag 1080
cccctagaga acctgggaat gaccgacatg ttcagacagt ttcaggctga cttcacgagt 1140
ctttcagacc aagagcctct ccacgtcgcg caggcgctgc agaaagtgaa gatcgaggtg 1200
aacgagagtg gcacggtggc ctcctcatcc acagctgtca tagtctcagc ccgcatggcc 1260
cccgaggaga tcatcatgga cagacccttc ctctttgtgg tccctcccca gaaacagtgt 1320
gcatgggtta ttttggagtg taggtgactt gtttactcat tgaagcagat ttctgcttcc 1380
ttttatttt ataggaatag aggaagaaag gtcagatgcg tgcccagctc ttcacccccc 1440
aatctcttgg tggggagggg tgtacctaaa tatttatcat atccttgccc ttgagtgctt 1500
gttagagaga aagagaacta ctaaggaaaa taatattatt taaactcgct cctagtgttt 1560
ctttgtggtc tgtgtcaccg tatctcagga agtccagcca cttgactggc acacacccct 1620
ccggacatcc agcgtgacgg agcccacact gccaccttgt ggccgcctga gaccctcgcg 1680
ccccccgcgc cccccgcgcc cctcttttttc cccttgatgg aaattgacca tacaatttca 1740
tcctccttca ggggatcaaa aggacggagt ggggggacag agactcagat gaggacagag 1800
tggtttccaa tgtgttcaat agatttagga gcagaaatgc aaggggctgc atgacctacc 1860
aggacagaac tttccccaat tacagggtga ctcacagccg cattggtgac tcacttcaat 1920
gtgtcatttc cggctgctgt gtgtgagcag tggacacgtg agggggggggt gggtgagaga 1980
gacaggcagc tcggattcaa ctaccttaga taatatttct gaaaacctac cagccagagg 2040
gtagggcaca aagatggatg taatgcactt tgggaggcca aggcgggagg attgcttgag 2100
cccaggagtt caagaccagc ctgggcaaca taccaagacc cccgtctctt taaaaatata 2160
tatattttaa atatacttaa atatatattt ctaatatctt taaatatata tatatatttt 2220
aaagaccaat ttatgggaga attgcacaca gatgtgaaat gaatgtaatc taatagaagc 2280
ctaatcagcc caccatgttc tccactgaaa aatcctcttt ctttggggtt tttctttctt 2340
tctttttttga ttttgcactg gacggtgacg tcagccatgt acaggatcca caggggtggt 2400
gtcaaatgct attgaaattg tgttgaattg tatgcttttt cacttttgat aaataaacat 2460
gtaaaaatgt ttcaaaaaaa taataaaata aataaatacg aagaatatgt caggacaaaa 2520
aa 2522

<210> 12
<211> 2272
<212> DNA
<213> HUMAN

<400> 12
ctcttccccg ctgagaggsa cagccastgc gwctccwccc tccagctcga cggcagccgc 60
ccggccgaca gccccgagac gacagcccgg cgcgtcccgg tccccacctc cgaccaccgc 120
cagcgctcca ggccccgcgc tccccgctcg ccgccaccgc gccctccgct ccgcccgcag 180
tgccaaccat gaccgccgcc agtatgggcc ccgtccgcgt cgccttcgtg gtcctcctcg 240
ccctctgcag ccggccggcc gtcggccaga actgcagcgg gccgtgccgg tgcccggacg 300
agccggcgcc gcgctgcccg gcgggcgtga gcctcgtgct ggacggctgc ggctgctgcc 360
gcgtctgcgc caagcagctg ggcgagctgt gcaccgagcg cgacccctgc gacccgcaca 420
agggcctctt ctgtgacttc ggctccccgg ccaaccgcaa gatcggcgtg tgcaccgcca 480
aagatggtgc tccctgcatc ttcggtggta cggtgtaccg cagcggagag tccttccaga 540
gcagctgcaa gtaccagtgc acgtgcctgg acggggcggt gggctgcatg cccctgtgca 600
gcatggacgt tcgtctgccc agccctgact gccccttacc gactggaaga cacgtttggc 660
ccagacccaa ctatgattag agccaactgc ctggtccaga ccacagagtg gagcgcctgt 720
tccaagacct gtgggatggg catctccacc cgggttacca atgacaacgc ctcctgcagg 780
ctagagaagc agagccgcct gtgcatggtc aggccttgcg aagtgacctg gaagagaaca 840
ttaagaaggg caaaaagtgc atccgtactc ccaaaatctc caagcctatc aagtttgagc 900
tttctggctg caccagcatg aagacatacc gagctaaatt ctgtggagta tgtaccgacg 960
gccgatgctg caccccccac agaaccacca ccctgccggt ggagttcaag tgccctgacg 1020
gcgaggtcat gaagaagaac atgatgttca tcaagacctg tgcctgccat tacaactgtc 1080
ccggagacaa tgacatcttt gaatcgctgt actacaggaa gatgtacgga gacatggcat 1140
gaagccagag agtgagagac attaactcat tagactggaa cttgaactga ttcacatctc 1200
atttttccgt aaaaatgatt tcagtagcac aagttattta aatctgtttt tctaactggg 1260
ggaaaagatt cccacccaat tcaaaacatt gtgccatgtc aaacaaatag tctatcaacc 1320
ccagacactg gtttgaagaa tgttaagact tgacagtgga actacattag tacacagcac 1380
cagaatgtat attaaggtgt ggctttagga gcagtgggag ggtaccagca gaaaggttag 1440
tatcatcaga tagcatctta tacgagtaat atgcctgcta tttgaagtgt aattgagaag 1500

gaaaatttta gcgtgctcac tgacctgcct gtagccccag tgacagctag gatgtgcatt 1560
ctccagccat caagagactg agtcaagttg ttccttaagt cagaacagca gactcagctc 1620
tgacattctg attcgaatga cactgttcag gaatcggaat cctgtcgatt agactggaca 1680
gcttgtggca agtgaatttg cctgtaacaa gccagatttt ttaaaattta tattgtaaat 1740
attgtgtgtg tgtgtgtgtg tgtatatata tatatatgta cagttatcta agttaattta 1800
aagttgtttg tgccttttta tttttgtttt taatgctttg atatttcaat gttagcctca 1860
atttctgaac accataggta gaatgtaaag cttgtctgat cgttcaaagc atgaaatgga 1920
tacttatatg gaaattctgc tcagatagaa tgacagtccg tcaaaacaga ttgtttgcaa 1980
aggggaggca tcagtgtcct tggcaggctg atttctaggt aggaaatgtg gtagcctcac 2040
ttttaatgaa caaatggcct ttattaaaaa ctgagtgact ctatatagct gatcagtttt 2100
ttcacctgga agcatttgtt tctactttga tatgactgtt tttcggacag tttatttgtt 2160
gagagtgtga ccaaaagtta catgtttgca cctttctagt tgaaaataaa gtgtatattt 2220
tttctataaa aaatgtcgac cggccgctaa tttagtagta gtaggcggca gc 2272

<210> 13
<211> 1153
<212> DNA
<213> HUMAN

<400> 13
tgaaatatgg agaatagtct tagatgtgtt tgggtaccca agctggcttt tgtactcttc 60
ggagcttcct tgctcagcgc gcatcttcaa gtaaccggtt ttcaaattaa agctttcaca 120
gcactgcgct tcctctcaga accttctgat gccgtcacaa tgcggggagg aaatgtcctc 180
ctcgactgct ccgcggagtc cgaccgagga gttccagtga tcaagtggaa gaaagatgca 240
ttcatctggc cttgggaatg gatgaaagga agcagcaact ttcaaatggg tctctgctga 300
tacaaaacat acttcattcc agacaccaca agccagatga gggactttac caatgtgagg 360
catctttagg agattctggc tcaattatta gtcggacagc aaaagttgca gtagcaggac 420
cactgaggtt cctttcacag acagaatctg tcacagcctt catgggagac acagtgctac 480
tcaagtgtga agtcattggg gagcccatgc caacaatcca ctggcagaag aaccaacaag 540
acctgactcc aatcccaggt gactcccgag tggtggtctt gccctctgga gcattgcaga 600
tcagccgact ccaaccgggg gacattggaa tttaccgatg ctcagctcga aatccagcca 660
gctcaagaac aggaaatgaa gcagaagtca gaattttatc agatccagga ctgcatagac 720
agctgtattt tctgcaaaga ccatccaatg tagtagccat tgaaggaaaa gatgctgtcc 780
tggaatgttg tgtttctggc tatcctccac caagttttac ctggttacga ggcgaggaag 840
tcatccaact caggtctaaa aagtattctt tattgggtgg aagcaacttg cttatctcca 900
atgtgacaga tgatgacagt ggaatgtata cctgtgttgt cacatataaa aatgagaata 960
ttagtgcctc tgcagagctc acagtcttgg taatcatcga taaagtattg gtggacacct 1020
tttgggtata aggattgcac taggttctgg gaagactaca gaaacatctg aaacctggct 1080
tcagtgggct tcgtataatt acaattcgta aaatcagttg gccwaatact gtctgaataa 1140
acttttwwtt aaa 1153

<210> 14
<211> 4459
<212> DNA
<213> HUMAN

<400> 14
tgaaatatgg agaatagtct tagatgtgtt tgggtaccca agctggcttt tgtactcttc 60
ggagcttcct tgctcagcgc gcatcttcaa gtaaccggtt ttcaaattaa agctttcaca 120
gcactgcgct tcctctcaga accttctgat gccgtcacaa tgcggggagg aaatgtcctc 180
ctcgactgct ccgcggagtc cgaccgagga gttccagtga tcaagtggaa gaaagatgca 240
ttcatctggc cttgggaatg gatgaaagga agcagcaact ttcaaatggg tctctgctga 300
tacaaaacat acttcattcc agacaccaca agccagatga gggactttac caatgtgagg 360
catctttagg agattctggc tcaattatta gtcggacagc aaaagttgca gtagcaggac 420
cactgaggtt cctttcacag acagaatctg tcacagcctt catgggagac acagtgctac 480
tcaagtgtga agtcattggg gagcccatgc caacaatcca ctggcagaag aaccaacaag 540
acctgactcc aatcccaggt gactcccgag tggtggtctt gccctctgga gcattgcaga 600
tcagccgact ccaaccgggg gacattggaa tttaccgatg ctcagctcga aatccagcca 660
gctcaagaac aggaaatgaa gcagaagtca gaattttatc agatccagga ctgcatagac 720

agctgtattt tctgcaaaga ccatccaatg tagtagccat tgaaggaaaa gatgctgtcc 780
tggaatgttg tgtttctggc tatcctccac caagttttac ctggttacga ggcgaggaag 840
tcatccaact caggtctaaa aagtattctt tattgggtgg aagcaacttg cttatctcca 900
atgtgacaga tgatgacagt ggaatgtata cctgtgttgt cacatataaa aatgagaata 960
ttagtgcctc tgcagagctc acagtcttgg ttccgccatg gtttttaaat catccttcca 1020
acctgtatgc ctatgaaagc atggatattg agtttgaatg tacagtctct ggaaagcctg 1080
tgcccactgt gaattggatg aagaatggag atgtggtcat tcctagtgat tattttcaga 1140
tagtgggagg aagcaactta cggatacttg gggtggtgaa gtcagatgaa ggcttttatc 1200
aatgtgtggc tgaaaatgag gctggaaatg cccagaccag tgcacagctc attgtcccta 1260
agcctgcaat cccaagctcc agtgtcctcc cttcggctcc cagagatgtg gtccctgtct 1320
tggtttccag ccgatttgtc cgtctcagct ggcgcccacc tgcagaagcg aaagggaaca 1380
ttcaaacttt cacggtcttt ttctccagag aaggtgacaa cagggaacga gcattgaata 1440
caacacagcc tgggtccctt cagctcactg tgggaaacct gaagccagaa gccatgtaca 1500
cctttcgagt tgtggcttac aatgaatggg gaccgggaga gagttctcaa cccatcaagg 1560
tggccacaca gcctgagttg caagttccag ggccagtaga aaacctgcaa gctgtatcta 1620
cctcacctac ctcaattctt attacctggg aacccccctgc ctatgcaaac ggtccagtcc 1680
aaggttacag attgttctgc actgaggtgt ccacaggaaa agaacagaat atagaggttg 1740
atggactatc ttataaactg gaaggcctga aaaaattcac cgaatatagt cttcgattct 1800
tagcttataa tcgctatggt ccgggcgtct ctactgatga tataacagtg gttacacttt 1860
ctgacgtgcc aagtgccccg cctcagaacg tctccctgga agtggtcaat tcaagaagta 1920
tcaaagttag ctggctgcct cctccatcag gaacacaaaa tggatttatt accggctata 1980
aaattcgaca cagaaagacg acccgcaggg gtgagatgga aacactggag ccaaacaacc 2040
tctggtacct attcacagga ctggagaaag gaagtcagta cagtttccag gtgtcagcca 2100
tgacagtcaa tggtactgga ccaccttcca actggtatac tgcagagact ccagagaatg 2160
atctagatga atctcaagtt cctgatcaac caagctctct tcatgtgagg ccccagacta 2220
actgcatcat catgagttgg actcctccct tgaacccaaa catcgtggtg cgaggttata 2280
ttatcggtta tggcgttggg agcccttacg ctgagacagt gcgtgtggac agcaagcagc 2340
gatattattc cattgagagg ttagagtcaa gttcccatta tgtaatctcc ctaaaagctt 2400
ttaacaatgc cggagaagga gttcctcttt atgaaagtgc caccaccagg tctataaccg 2460
atcccactga cccagttgat tattatcctt tgcttgatga tttccccacc tcggtcccag 2520
atctctccac ccccatgctc ccaccagtag gtgtacaggc tgtggctctt acccatgatg 2580
ctgtgagggt cagctgggca gacaactctg tccctaagaa ccaaaagacg tctgaggtgc 2640
gactttacac cgtccggtgg agaaccagct tttctgcaag tgcaaaatac aagtcagaag 2700
acacaacatc tctaagttac acagcaacag gcctcaaacc aaacacaatg tatgaattct 2760
cggtcatggt aacaaaaaac agaaggtcca gtacttggag catgactgca catgccacca 2820
cgtatgaagc agccccacc tctgctccca aggactttac agtcattact agggaaggga 2880
agcctcgtgc cgtcattgtg agttggcagc ctcccttgga agccaatggg aaaattactg 2940
cttacatctt attttatacc ttggacaaga acatcccaat tgatgactgg attatggaaa 3000
caatcagtgg tgataggctt actcatcaaa tcatggatct caaccttgat actatgtatt 3060
actttcgaat tcaagcacga aattcaaaag gagtggggcc actctctgat cccatcctct 3120
tcaggactct gaaagtggaa caccctgaca aaatggctaa tgaccaaggt cgtcatggag 3180
atggaggtta ttggccagtt gatactaatt tgattgatag aagcacccta aatgagccgc 3240
caattggaca aatgcacccc ccgcatggca gtgtcactcc tcagaagaac agcaacctgc 3300
ttgtgatcat tgtggtcacc gttggtgtca tcacagtgct ggtagtggtc atcgtggctg 3360
tgatttgcac ccgacgctct tcagcccagc agagaaagaa acgggccacc cacagtgctg 3420
gcaaaaggaa gggcagccag aaggacctcc gacccccctga tctttggatc catcatgaag 3480
aaatggagat gaaaaatatt gaaaagccat ctggcactga ccctgcagga agggactctc 3540
ccatccaaag ttgccaagac ctcacaccag tcagccacag ccagtcagaa acccaactgg 3600
gaagcaaaag cacctctcat tcaggtcaag acactgagga agcagggagc tctatgtcca 3660
ctctggagag gtcgctggct gcacgccgag ccccccgggc caagctcatg attcccatgg 3720
atgcccagtc caacaatcct gctgtcgtga gcgccatccc ggtgccaacg ctagaaagtg 3780
cccagtaccc aggaatcctc ccgtctccca cctgtggata tccccacccg cagttcactc 3840
tccggcctgt gccattccca acactctcag tggaccgagg tttcggagca ggaagaagtc 3900
agtcagtgag tgaaggacca actacccaac aaccacctat gctgccccca tctcagcctg 3960
agcattctag cagcgaggag gcaccaagca gaaccatccc cacagcttgt gttcgaccaa 4020
ctcacccact ccgcagcttt gctaatcctt tgctacctcc accaatgagt gcaatagaac 4080
cgaaagtccc ttacacacca cttttgtctc agccagggcc cactcttcct aagacccatg 4140
tgaaaacagc ctcccttggg ttggctggaa aagcaagatc ccctttgctt cctgtgtctg 4200
tgccaacagc ccctgaagtg tctgaggaga gccacaaacc aacagaggat tcagccaatg 4260
taagtgcatc tttaaaattc atgcttcatc aagggacaga ttgactggcg ctgtgtaata 4320
gcatctaaac caatgagtac tgtccattca ccccaggcca ccccaggcca ccccagtcca 4380

ttgttgttct tattgttgtt tgttcctttg ttttcctgca cttaaatatg aagaggaaat 4440
taaatctatg atattaaaa 4459

<210> 15
<211> 3744
<212> DNA
<213> HUMAN

<400> 15
ccgagtgaca aggaggtggg agagggtagc agcatgggct acgcggttgg ctgccctcag 60
tccccctgct gctgaagctg ccctgcccat gcccacccag gccgtggggc caggggcctg 120
ccagggctag gagtgggcct gccgttcatg ggtctctagg gatttccgar atgcctggga 180
agagaggctt gggctggtgg tgggcccggc tgccccttg cctgctcctc agcctttacg 240
gcccctggat gccttcctcc ctgggaaagc ccaaaggcca ccctcacatg aattccatcc 300
gcatagatgg ggacatcaca ctgggaggcc tgttcccggt gcatggccgg ggctcagagg 360
gcaagccctg tggagaactt aagaaggaaa agggcatcca ccggctggag gccatgctgt 420
tcgccctgga tcgcatcaac aacgacccgg acctgctgcc taacatcacg ctgggcgccc 480
gcattctgga cacctgctcc agggacaccc atgccctcga gcagtcgctg acctttgtgc 540
aggcgctcat cgagaaggat ggcacagagg tccgctgtgg cagtggcggc ccacccatca 600
tcaccaagcc tgaacgtgtg gtgggtgtca tcggtgcttc agggagctcg gtctccatca 660
tggtggccaa catccttcgc ctcttcaaga taccccagat cagctacgcc tccacagcgc 720
cagacctgag tgacaacagc cgctacgact tcttctcccg cgtggtgccc tcggacacgt 780
accaggccca ggccatggtg gacatcgtcc gtgccctcaa gtggaactat gtgtccacag 840
tggcctcgga gggcagctat ggtgagagcg gtgtggaggc cttcatccag aagtcccgtg 900
aggacggggg cgtgtgcatc gcccagtcgg tgaagatacc acgggagccc aaggcaggcg 960
agttcgacaa gatcatccgc cgcctcctgg agacttcgaa cgccagggca gtcatcatct 1020
ttgccaacga ggatgacatc aggcgtgtgc tggaggcagc acgaagggcc aaccagacag 1080
gccatttctt ctggatgggc tctgacagct ggggctccaa gattgcacct gtgctgcacc 1140
tggaggaggt ggctgagggt gctgtcacga tcctccccaa gaggatgtcc gtacgagacc 1200
gtgagcgaat tgggcaggat tcagcttatg agcaggaggg gaaggtgcag tttgtgatcg 1260
atgccgtgta cgccatgggc cacgcgctgc acgccatgca ccgtgacctg tgtcccggcc 1320
gcgtggggct ctgcccgcgc atggaccctg tagatggcac ccagctgctt aagtacatcc 1380
gaaacgtcaa cttctcaggc atcgcaggga accctgtgac cttcaatgag aatggagatg 1440
cgcctgggcg ctatgacatc taccaatacc agctgcgcaa cgattctgcc gagtacaagg 1500
tcattggctc ctggactgac cacctgcacc ttagaataga gcggatgcac tggccgggga 1560
gcgggcagca gctgccccgc tccatctgca gcctgccctg ccaaccgggt gagcggaaga 1620
agacagtgaa gggcatgcct tgctgctggc actgcgagcc ttgcacaggg taccagtacc 1680
aggtggaccg ctacacctgt aagacgtgtc cctatgacat gcggcccaca gagaaccgca 1740
cgggctgccg gcccatcccc atcatcaagc ttgagtgggg ctcgccctgg gccgtgctgc 1800
ccctcttcct ggccgtggtg ggcatcgctg ccacgttgtt cgtggtgatc acctttgtgc 1860
gctacaacga cacgcccatc gtcaaggcct cgggccgtga actgagctac gtgctgctgg 1920
caggcatctt cctgtgctat gccaccacct tcctcatgat cgctgagccc gaccttggca 1980
cctgctcgct gcgccgaatc ttcctgggac tagggatgag catcagctat gcagccctgc 2040
tcaccaagac caaccgcatc taccgcatct tcgagcaggg caagcgctcg gtcagtgccc 2100
cacgcttcat cagccccgcc tcacagctgg ccatcacctt cagcctcatc tcgctgcagc 2160
tgctgggcat ctgtgtgtgg tttgtggtgg acccctccca ctcggtggtg gacttccagg 2220
accagcggac actcgacccc cgcttcgcca ggggtgtgct caagtgtgac atctcggacc 2280
tgtcgctcat ctgcctgctg ggctacagca tgctgctcat ggtcacgtgc accgtgtatg 2340
ccatcaagac acgcggcgtg cccgagacct tcaatgaggc caagcccatt ggcttcacca 2400
tgtacaccac ttgcatcgtc tggctggcct tcatccccat cttctttggc acctcgcagt 2460
cggccgacaa gctgtacatc cagacgacga cgctgacggt ctcggtgagt ctgagcgcct 2520
cggtgtccct gggaatgctc tacatgccca aagtctacat catcctcttc cacccggagc 2580
agaacgtgcc caagcgcaag cgcagcctca aagccgtcgt tacggcggcc accatgtcca 2640
acaagttcac gcagaagggc aacttccggc ccaacggaga ggccaagtct gagctctgcg 2700
agaaccttga ggccccagcg ctggccacca aacagactta cgtcacttac accaaccatg 2760
caatctagcg agtccatgga gctgagcagc aggaggagga gccgtgaccc tgtggaaggt 2820
gcgtcgggcc agggccacac ccaagggccc agctgtcttg cctgcccgtg ggcacccacg 2880
gacgtggctt ggtgctgagg atagcagagc ccccagccat cactgctggc agcctgggca 2940
aaccgggtga gcaacaggag gacgaggggc cggggcggtg ccaggctacc acaagaacct 3000
gcgtcttgga ccattgcccc tcccggcccc aaaccacagg ggctcaggtc gtgtgggccc 3060

cagtgctaga tctctccctc ccttcgtctc tgtctgtgct gttggcgacc cctctgtctg 3120
tctccagccc tgtctttctg ttctcttatc tctttgtttc acctttttccc tctctggcgt 3180
ccccgyctgc ttgtactctt ggccttttct gtgtctcctt tctggctctt gcctccgcct 3240
ctctctctca tcctctttgt cctcagctcc tcctgctttc ttgggtccca ccagtgtcac 3300
ttttctgccg ttttctttcc tgttctcctc tgcttcattc tcgtccagcc attgctcccc 3360
tctccctgcc acccttcccc agttcaccaa accttacatg ttgcaaaaga gaaaaaagga 3420
aaaaaaatca aaacacaaaa aagccaaaac gaaaacaaat ctcgagtgtg ttgccaagtg 3480
ctgcgtcctc ctggtggcct ctgtgtgtgt ccctgtggcc cgcagcctgc ccgcctgccc 3540
cgcccatctg ccgtgtgtct tgcccgcctg ccccgcccgt ctgccgtctg tcttgcccgc 3600
ctgcccgcct gcccctcctg ccgaccacac ggagttcagt gcctgggtgt ttggtgatgg 3660
ttattgacga caatgtgtag cgcatgattg tttttatacc aagaacattt ctaataaaaa 3720
taaacacatg gttttgcaaa aaaa 3744

<210> 16
<211> 2899
<212> DNA
<213> HUMAN

<400> 16
cgcgggagcc aacttcaggc tgctcagagg aagcccgtgc agtcagtcac ctgggtgcaa 60
gagcgttgct gcctcgggct ctcccgctgc agggagagcg gcactcgctg gcctggatgt 120
ggttggattt agggggggctc cgcagcaggg gtttcgtggc ggtggcaagc gctgcaacag 180
gtagacggcg agagacggac cccggccgag gcagggatgg agaccaaagg ctaccacagt 240
ctccctgaag gtctagatat ggaaagacgg tggggtcaag tttctcaggc tgtggagcgt 300
tcttccctgg gacctacaga gaggaccgat gagaataact acatggagat tgtcaacgta 360
agctgtgttt ccggtgctat tccaaacaac agtactcaag gaagcagcaa agaaaaacaa 420
gaactactcc cttgccttca gcaagacaat aatcggcctg ggattttaac atctgatatt 480
aaaactgagc tggaatctaa ggaactttca gcaactgtag ctgagtccat gggtttatat 540
atggattctg taagagatgc tgactattcc tatgagcagc agaaccaaca aggaagcatg 600
agtccagcta agatttatca gaatgttgaa cagctggtga aattttacaa aggaaatggc 660
catcgtcctt ccactctaag ttgtgtgaac acgcccttga gatcatttat gtctgactct 720
gggagctccg tgaatggtgg cgtcatgcgc gccattgtta aaagccctat catgtgtcat 780
gagaaaagcc cgtctgtttg cagccctctg aacatgacat cttcggtttg cagccctgct 840
ggaatcaact ctgtgtcctc caccacagcc agctttggca gttttccagt gcacagccca 900
atcacccagg gaactcctct gacatgctcc cctaatgctg aaaatcgagg ctccaggtcg 960
cacagccctg cacatgctag caatgtgggc tctcctctct caagtccgtt aagtagcatg 1020
aaatcctcaa tttccagccc tccaagtcac tgcagtgtaa aatctccagt ctccagtccc 1080
aataatgtca ctctgagatc ctctgtgtct agccctgcaa atattaacaa ctcaaggtgc 1140
tctgtttcca gcccttcgaa cactaataac agatccacgc tttccagtcc ggcagccagt 1200
actgtgggat ctatctgtag ccctgtaaac aatgccttca gctacactgc ttctggcacc 1260
tctgctggat ccagtacatt gcgggatgtg gttcccagtc cagacacgca ggagaaaggt 1320
gctcaagagg tccctttttcc taagactgag gaagtagaga gtgccatctc aaatggtgtg 1380
actggccagc ttaatattgt ccagtacata aaaccagaac cagatggagc ttttagcagc 1440
tcatgtctag gaggaaatag caaaataaat tcggattctt cattctcagt accaataaag 1500
caagaatcaa ccaagcattc atgttcaggc acctctttta aagggaatcc aacagtaaac 1560
ccgtttccat ttatggatgg ctcgtatttt tcctttatgg atgataaaga ctattattcc 1620
ctatcaggaa ttttaggacc acctgtgccc ggctttgatg gtaactgtga aggcagcgga 1680
ttcccagtgg gtattaaaca agaaccagat gacgggagct attacccaga ggccagcatc 1740
ccttcctctg ctattgttgg ggtgaattca ggtggacagt ccttccacta caggattggt 1800
gctcaaggta caatatcttt atcacgatcg gctagagacc aatctttcca acacctgagt 1860
tcctttcctc ctgtcaatac tttagtggag tcatggaaat cacacggcga cctgtcgtct 1920
agaagaagtg atgggtatcc ggtcttagaa tacattccag aaaatgtatc aagctctact 1980
ttacgaagtg tttctactgg atcttcaaga ccttcaaaaa tatgtttggt gtgtggggat 2040
gaggcttcag gatgccatta tggggtagtc acctgtggca gctgcaaagt tttcttcaaa 2100
agagcagtgg aagggcaaca caactattta tgtgctggaa gaaatgattg catcattgat 2160
aagattcgac gaaagaattg tcctgcttgc agacttcaga aatgtcttca agctggaatg 2220
aatttaggag cacgaaagtc aaagaagttg ggaaagttaa aagggattca cgaggagcag 2280
ccacagcagc agcagccccc acccccaccc ccacccccgc aaagcccaga ggaagggaca 2340
acgtacatcg ctcctgcaaa agaaccctcg gtcaacacag cactggttcc tcagctctcc 2400
acaatctcac gagcgctcac accttccccc gttatggtcc ttgaaaacat tgaacctgaa 2460

attgtatatg caggctatga cagctcaaaa ccagatacag ccgaaaatct gctctccacg 2520
ctcaaccgct tagcaggcaa acagatgatc caagtcgtga agtgggcaaa ggtacttcca 2580
ggatttaaaa acttgcctct tgaggaccaa attaccctaa tccagtattc ttggatgtgt 2640
ctatcatcat ttgccttgag ctggagatcg tacaaacata cgaacagcca atttctctat 2700
tttgcaccag acctagtctt taatgaactg ttggcgagag tgagagaggg ttgaagtctt 2760
agtggcactt cacatgcctc tgtggattct ctgagtgtct ctctgggatg aaggtgactt 2820
ccacccacac catcttgcgt ctgaagcact agttgtgtga gggcagctcc attagacgag 2880
tacacaactc agttgcaaa 2899

<210> 17
<211> 2780
<212> DNA
<213> HUMAN

<400> 17
cgcgggagcc aacttcaggc tgctcagagg aagcccgtgc agtcagtcac ctgggtgcaa 60
gagcgttgct gcctcgggct ctcccgctgc agggagagcg gcactcgctg gcctggatgt 120
ggttggattt agggggggctc cgcagcaggg gtttcgtggc ggtggcaagc gctgcaacag 180
gtagacggcg agagacggac cccggccgag gcagggatgg agaccaaagg ctaccacagt 240
ctccctgaag gtctagatat ggaaagacgg tggggtcaag tttctcaggc tgtggagcgt 300
tcttccctgg gacctacaga gaggaccgat gagaataact acatggagat tgtcaacgta 360
agctgtgttt ccggtgctat tccaaacaac agtactcaag gaagcagcaa agaaaaacaa 420
gaactactcc cttgccttca gcaagacaat aatcggcctg ggattttaac atctgatatt 480
aaaactgagc tggaatctaa ggaactttca gcaactgtag ctgagtccat gggtttatat 540
atggattctg taagagatgc tgactattcc tatgagcagc agaaccaaca aggaagcatg 600
agtccagcta agatttatca gaatgttgaa cagctggtga aattttacaa aggaaatggc 660
catcgtcctt ccactctaag ttgtgtgaac acgcccttga gatcatttat gtctgactct 720
gggagctccg tgaatggtgg cgtcatgcgc gccattgtta aaagccctat catgtgtcat 780
gagaaaagcc cgtctgtttg cagccctctg aacatgacat cttcggtttg cagccctgct 840
ggaatcaact ctgtgtcctc caccacagcc agctttggca gttttccagt gcacagccca 900
atcacccagg gaactcctct gacatgctcc cctaatgctg aaaatcgagg ctccaggtcg 960
cacagccctg cacatgctag caatgtgggc tctcctctct caagtccgtt aagtagcatg 1020
aaatcctcaa tttccagccc tccaagtcac tgcagtgtaa aatctccagt ctccagtccc 1080
aataatgtca ctctgagatc ctctgtgtct agccctgcaa atattaacaa ctcaaggtgc 1140
tctgtttcca gcccttcgaa cactaataac agatccacgc tttccagtcc ggcagccagt 1200
actgtgggat ctatctgtag ccctgtaaac aatgccttca gctacactgc ttctggcacc 1260
tctgctggat ccagtacatt gcgggatgtg gttcccagtc cagacacgca ggagaaaggt 1320
gctcaagagg tccctttttcc taagactgag gaagtagaga gtgccatctc aaatggtgtg 1380
actggccagc ttaatattgt ccagtacata aaaccagaac cagatggagc ttttagcagc 1440
tcatgtctag gaggaaatag caaaataaat tcggattctt cattctcagt accaataaag 1500
caagaatcaa ccaagcattc atgttcaggc acctctttta aagggaatcc aacagtaaac 1560
ccgtttccat ttatggatgg ctcgtatttt tcctttatgg atgataaaga ctattattcc 1620
ctatcaggaa ttttaggacc acctgtgccc ggctttgatg gtaactgtga aggcagcgga 1680
ttcccagtgg gtattaaaca agaaccagat gacgggagct attacccaga ggccagcatc 1740
ccttcctctg ctattgttgg ggtgaattca ggtggacagt ccttccacta caggattggt 1800
gctcaaggta caatatcttt atcacgatcg gctagagacc aatctttcca acacctgagt 1860
tcctttcctc ctgtcaatac tttagtggag tcatggaaat cacacggcga cctgtcgtct 1920
agaagaagtg atgggtatcc ggtcttagaa tacattccag aaaatgtatc aagctctact 1980
ttacgaagtg tttctactgg atcttcaaga ccttcaaaaa tatgtttggt gtgtggggat 2040
gaggcttcag gatgccatta tggggtagtc acctgtggca gctgcaaagt tttcttcaaa 2100
agagcagtgg aagggcaaca caactattta tgtgctggaa gaaatgattg catcattgat 2160
aagattcgac gaaagaattg tcctgcttgc agacttcaga aatgtcttca agctggaatg 2220
aatttaggag cacgaaagtc aaagaagttg ggaaagttaa aagggattca cgaggagcag 2280
ccacagcagc agcagccccc acccccaccc ccacccccgc aaagcccaga ggaagggaca 2340
acgtacatcg ctcctgcaaa agaaccctcg gtcaacacag cactggttcc tcagctctcc 2400
acaatctcac gagcgctcac accttccccc gttatggtcc ttgaaaacat tgaacctgaa 2460
attgtatatg caggctatga cagctcaaaa ccagatacag ccgaaaatct gctctccacg 2520
ctcaaccgct tagcaggcaa acagatgatc caagtcgtga agtgggcaaa ggtacttcca 2580
ggatttaaaa acttgcctct tgaggaccaa attaccctaa tccagtattc ttggatgtgt 2640
ctatcatcat ttgccttgag ctggagatcg tacaaacata cgaacagcca atttctctat 2700

tttgcaccag acctagtctt taatgaataa gtacctatta attagccttc ataaaataaa 2760
ctgcagagtt atgcgttaaa 2780

<210> 18
<211> 481
<212> PRT
<213> HUMAN

<400> 18

```
Met Arg Glu Asn Met Ala Arg Gly Pro Cys Asn Ala Pro Arg Trp Val
  1               5                  10                  15

Ser Leu Met Val Leu Val Ala Ile Gly Thr Ala Val Thr Ala Ala Val
             20                  25                  30

Asn Pro Gly Val Val Val Arg Ile Ser Gln Lys Gly Leu Asp Tyr Ala
         35                  40                  45

Ser Gln Gln Gly Thr Ala Ala Leu Gln Lys Glu Leu Lys Arg Ile Lys
     50                  55                  60

Ile Pro Asp Tyr Ser Asp Ser Phe Lys Ile Lys His Leu Gly Lys Gly
 65                  70                  75                  80

His Tyr Ser Phe Tyr Ser Met Asp Ile Arg Glu Phe Gln Leu Pro Ser
                85                  90                  95

Ser Gln Ile Ser Met Val Pro Asn Val Gly Leu Lys Phe Ser Ile Ser
            100                 105                 110

Asn Ala Asn Ile Lys Ile Ser Gly Lys Trp Lys Ala Gln Lys Arg Phe
        115                 120                 125

Leu Lys Met Ser Gly Asn Phe Asp Leu Ser Ile Glu Gly Met Ser Ile
    130                 135                 140

Ser Ala Asp Leu Lys Leu Gly Ser Asn Pro Thr Ser Gly Lys Pro Thr
145                 150                 155                 160

Ile Thr Cys Ser Ser Cys Ser Ser His Ile Asn Ser Val His Val His
                165                 170                 175

Ile Ser Lys Ser Lys Val Gly Trp Leu Ile Gln Leu Phe His Lys Lys
            180                 185                 190

Ile Glu Ser Ala Leu Arg Asn Lys Met Asn Ser Gln Val Cys Glu Lys
        195                 200                 205

Val Thr Asn Ser Val Ser Ser Lys Leu Gln Pro Tyr Phe Gln Thr Leu
    210                 215                 220

Pro Val Met Thr Lys Ile Asp Ser Val Ala Gly Ile Asn Tyr Gly Leu
225                 230                 235                 240

Val Ala Pro Pro Ala Thr Thr Ala Glu Thr Leu Asp Val Gln Met Lys
                245                 250                 255

Gly Glu Phe Tyr Ser Glu Asn His His Asn Pro Pro Pro Phe Ala Pro
            260                 265                 270
```

```
Pro Val Met Glu Phe Pro Ala Ala His Asp Arg Met Val Tyr Leu Gly
        275                 280                 285

Leu Ser Asp Tyr Phe Phe Asn Thr Ala Gly Leu Val Tyr Gln Glu Ala
    290                 295                 300

Gly Val Leu Lys Met Thr Leu Arg Asp Asp Met Ile Pro Lys Glu Ser
305                 310                 315                 320

Lys Phe Arg Leu Thr Thr Lys Phe Phe Gly Thr Phe Leu Pro Glu Val
                325                 330                 335

Ala Lys Lys Phe Pro Asn Met Lys Ile Gln Ile His Val Ser Ala Ser
            340                 345                 350

Thr Pro Pro His Leu Ser Val Gln Pro Thr Gly Leu Thr Phe Tyr Pro
        355                 360                 365

Ala Val Asp Val Gln Ala Phe Ala Val Leu Pro Asn Ser Ser Leu Ala
    370                 375                 380

Ser Leu Phe Leu Ile Gly Met Gly Lys Gln Phe Leu Gly Trp Thr Asp
385                 390                 395                 400

Glu Glu Pro Gln Thr Val Pro Thr Ala Leu Ser Leu Glu Ser Gly Asp
                405                 410                 415

His Val Asn Pro Val Trp Ile Gln Thr Trp Thr Val Ser Leu Arg Ser
            420                 425                 430

Leu Arg Leu Glu Ser Leu Tyr Ser Met Val Pro Thr Pro Gly Gly Ile
        435                 440                 445

His Ser Pro Ser His Ser Leu Val Arg Leu Phe Thr Tyr Ser Phe Asn
    450                 455                 460

Tyr Ser Phe Ser Gln Phe Leu Ile His Ser Xaa Ile His Ser Met Leu
465                 470                 475                 480

Ala
```

<210> 19
<211> 628
<212> PRT
<213> HUMAN

<400> 19

```
Met Gly Arg Leu Gln Leu Val Val Leu Gly Leu Thr Cys Cys Trp Ala
  1               5                  10                  15

Val Ala Ser Ala Ala Lys Leu Gly Ala Val Tyr Thr Glu Gly Gly Phe
            20                  25                  30

Val Glu Gly Val Asn Lys Lys Leu Gly Leu Leu Gly Asp Ser Val Asp
        35                  40                  45

Ile Phe Lys Gly Ile Pro Phe Ala Ala Pro Thr Lys Ala Leu Glu Asn
    50                  55                  60
```

```
Pro Gln Pro His Pro Gly Trp Gln Gly Thr Leu Lys Ala Lys Asn Phe
 65                  70                  75                  80

Lys Lys Arg Cys Leu Gln Ala Thr Ile Thr Gln Asp Ser Thr Tyr Gly
                85                  90                  95

Asp Glu Asp Cys Leu Tyr Leu Asn Ile Trp Val Pro Gln Gly Arg Lys
            100                 105                 110

Gln Val Ser Arg Asp Leu Pro Val Met Ile Trp Ile Tyr Gly Gly Ala
        115                 120                 125

Phe Leu Met Gly Ser Gly His Gly Ala Asn Phe Leu Asn Asn Tyr Leu
    130                 135                 140

Tyr Asp Gly Glu Glu Ile Ala Thr Arg Gly Asn Val Ile Val Val Thr
145                 150                 155                 160

Phe Asn Tyr Arg Val Gly Pro Leu Gly Phe Leu Ser Thr Gly Asp Ala
                165                 170                 175

Asn Leu Pro Gly Asn Tyr Gly Leu Arg Asp Gln His Met Ala Ile Ala
            180                 185                 190

Trp Val Lys Arg Asn Ile Ala Ala Phe Gly Gly Asp Pro Asn Asn Ile
        195                 200                 205

Thr Leu Phe Gly Glu Ser Ala Gly Gly Ala Ser Val Ser Leu Gln Thr
    210                 215                 220

Leu Ser Pro Tyr Asn Lys Gly Leu Ile Arg Arg Ala Ile Ser Gln Ser
225                 230                 235                 240

Gly Val Ala Leu Ser Pro Trp Val Ile Gln Lys Asn Pro Leu Phe Trp
                245                 250                 255

Ala Lys Lys Val Ala Glu Lys Val Gly Cys Pro Val Gly Asp Ala Ala
            260                 265                 270

Arg Met Ala Gln Cys Leu Lys Val Thr Asp Pro Arg Ala Leu Thr Leu
        275                 280                 285

Ala Tyr Lys Val Pro Leu Ala Gly Leu Glu Tyr Pro Met Leu His Tyr
    290                 295                 300

Val Gly Phe Val Pro Val Ile Asp Gly Asp Phe Ile Pro Ala Asp Pro
305                 310                 315                 320

Ile Asn Leu Tyr Ala Asn Ala Ala Asp Ile Asp Tyr Ile Ala Gly Thr
                325                 330                 335

Asn Asn Met Asp Gly His Ile Phe Ala Ser Ile Asp Met Pro Ala Ile
            340                 345                 350

Asn Lys Gly Asn Lys Lys Val Thr Glu Glu Asp Phe Tyr Lys Leu Val
        355                 360                 365

Ser Glu Phe Thr Ile Thr Lys Gly Leu Arg Gly Ala Lys Thr Thr Phe
    370                 375                 380
```

```
Asp Val Tyr Thr Glu Ser Trp Ala Gln Asp Pro Ser Gln Glu Asn Lys
385                 390                 395                 400

Lys Lys Thr Val Val Asp Phe Glu Thr Asp Val Leu Phe Leu Val Pro
                405                 410                 415

Thr Glu Ile Ala Leu Ala Gln His Arg Ala Asn Ala Lys Ser Ala Lys
            420                 425                 430

Thr Tyr Ala Tyr Leu Phe Ser His Pro Ser Arg Met Pro Val Tyr Pro
        435                 440                 445

Lys Trp Val Gly Ala Asp His Ala Asp Asp Ile Gln Tyr Val Phe Gly
    450                 455                 460

Lys Pro Phe Ala Thr Pro Thr Gly Tyr Arg Pro Gln Asp Arg Thr Val
465                 470                 475                 480

Ser Lys Ala Met Ile Ala Tyr Trp Thr Asn Phe Ala Lys Thr Gly Asp
                485                 490                 495

Pro Asn Met Gly Asp Ser Ala Val Pro Thr His Trp Glu Pro Tyr Thr
            500                 505                 510

Thr Glu Asn Ser Gly Tyr Leu Glu Ile Thr Lys Lys Met Gly Ser Ser
        515                 520                 525

Ser Met Lys Arg Ser Leu Arg Thr Asn Phe Leu Arg Tyr Trp Thr Leu
    530                 535                 540

Thr Tyr Leu Ala Leu Pro Thr Val Thr Asp Gln Glu Ala Thr Pro Val
545                 550                 555                 560

Pro Pro Thr Gly Asp Ser Glu Ala Thr Pro Val Pro Pro Thr Gly Asp
                565                 570                 575

Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro
            580                 585                 590

Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly
        595                 600                 605

Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Arg
    610                 615                 620

Ala His Leu Gly
625
```

<210> 20
<211> 641
<212> PRT
<213> HUMAN

<400> 20

```
Leu Leu Leu Leu Gly Phe Leu Leu Val Ser Leu Glu Ser Thr Leu Ser
  1               5                  10                  15

Ile Pro Pro Trp Glu Ala Pro Lys Glu His Lys Tyr Lys Ala Glu Glu
            20                  25                  30
```

```
His Thr Val Val Leu Thr Val Thr Gly Glu Pro Cys His Phe Pro Phe
         35                  40                  45

Gln Tyr His Arg Gln Leu Tyr His Lys Cys Thr His Lys Gly Arg Pro
     50                  55                  60

Gly Pro Gln Pro Trp Cys Ala Thr Thr Pro Asn Phe Asp Gln Asp Gln
 65                  70                  75                  80

Arg Trp Gly Tyr Cys Leu Glu Pro Lys Lys Val Lys Asp His Cys Ser
                 85                  90                  95

Lys His Ser Pro Cys Gln Lys Gly Gly Thr Cys Val Asn Met Pro Ser
            100                 105                 110

Gly Pro His Cys Leu Cys Pro Gln His Leu Thr Gly Asn His Cys Gln
        115                 120                 125

Lys Glu Lys Cys Phe Glu Pro Gln Leu Leu Arg Phe Phe His Lys Asn
    130                 135                 140

Glu Ile Trp Tyr Arg Thr Glu Gln Ala Ala Val Ala Arg Cys Gln Cys
145                 150                 155                 160

Lys Gly Pro Asp Ala His Cys Gln Arg Leu Ala Ser Gln Ala Cys Arg
                165                 170                 175

Thr Asn Pro Cys Leu His Gly Gly Arg Cys Leu Glu Val Glu Gly His
            180                 185                 190

Arg Leu Cys His Cys Pro Val Gly Tyr Thr Gly Pro Phe Cys Asp Val
        195                 200                 205

Asp Thr Lys Ala Ser Cys Tyr Asp Gly Arg Gly Leu Ser Tyr Arg Gly
    210                 215                 220

Leu Ala Arg Thr Thr Leu Ser Gly Ala Pro Cys Gln Pro Trp Ala Ser
225                 230                 235                 240

Glu Ala Thr Tyr Arg Asn Val Thr Ala Glu Gln Ala Arg Asn Trp Gly
                245                 250                 255

Leu Gly Gly His Ala Phe Cys Arg Asn Pro Asp Asn Asp Ile Arg Pro
            260                 265                 270

Trp Cys Phe Val Leu Asn Arg Asp Arg Leu Ser Trp Glu Tyr Cys Asp
        275                 280                 285

Leu Ala Gln Cys Gln Thr Pro Thr Gln Ala Ala Pro Pro Thr Pro Val
    290                 295                 300

Ser Pro Arg Leu His Val Pro Leu Met Pro Ala Gln Pro Ala Pro Pro
305                 310                 315                 320

Lys Pro Gln Pro Thr Thr Arg Thr Pro Pro Gln Ser Gln Thr Pro Gly
                325                 330                 335

Ala Leu Pro Ala Lys Arg Glu Gln Pro Pro Ser Leu Thr Arg Asn Gly
            340                 345                 350
```

```
Pro Leu Ser Cys Gly Gln Arg Leu Arg Lys Ser Leu Ser Ser Met Thr
        355                 360                 365

Arg Val Val Gly Gly Leu Val Ala Leu Arg Gly Ala His Pro Tyr Ile
    370                 375                 380

Ala Ala Leu Tyr Trp Gly His Ser Phe Cys Ala Gly Ser Leu Ile Ala
385                 390                 395                 400

Pro Cys Trp Val Leu Thr Ala Ala His Cys Leu Gln Asp Arg Pro Ala
                405                 410                 415

Pro Glu Asp Leu Thr Val Val Leu Gly Gln Glu Arg Arg Asn His Ser
            420                 425                 430

Cys Glu Pro Cys Gln Thr Leu Ala Val Arg Ser Tyr Arg Leu His Glu
        435                 440                 445

Ala Phe Ser Pro Val Ser Tyr Gln His Asp Leu Ala Leu Leu Arg Leu
    450                 455                 460

Gln Glu Asp Ala Asp Gly Ser Cys Ala Leu Leu Ser Pro Tyr Val Gln
465                 470                 475                 480

Pro Val Cys Leu Pro Ser Gly Ala Ala Arg Pro Ser Glu Thr Thr Leu
                485                 490                 495

Cys Gln Val Ala Gly Trp Gly His Gln Phe Glu Gly Ala Glu Glu Tyr
            500                 505                 510

Ala Ser Phe Leu Gln Glu Ala Gln Val Pro Phe Leu Ser Leu Glu Arg
        515                 520                 525

Cys Ser Ala Pro Asp Val His Gly Ser Ser Ile Leu Pro Gly Met Leu
    530                 535                 540

Cys Ala Gly Phe Leu Glu Gly Gly Thr Asp Ala Cys Ala Gly Glu Leu
545                 550                 555                 560

Leu Ala Gly Trp Arg Pro Ser Pro Arg Pro Ser Ala Xaa Ser Gln Val
                565                 570                 575

His Ser Ala Asp Cys Val Phe Pro Thr Gln Gly Asp Ser Gly Gly Pro
            580                 585                 590

Leu Val Cys Glu Asp Gln Ala Ala Glu Arg Arg Leu Thr Leu Gln Gly
        595                 600                 605

Ile Ile Ser Trp Gly Ser Gly Cys Gly Asp Arg Asn Lys Pro Gly Val
    610                 615                 620

Tyr Thr Asp Val Ala Tyr Tyr Leu Ala Trp Ile Arg Glu His Thr Val
625                 630                 635                 640

Ser
```

<210> 21
<211> 164

<212> PRT
<213> HUMAN

<400> 21
Met Ala Pro Phe Glu Pro Leu Ala Ser Gly Ile Leu Leu Leu Leu Trp
1 5 10 15

Leu Ile Ala Pro Ser Arg Ala Cys Thr Cys Val Pro Pro His Pro Gln
20 25 30

Thr Ala Phe Cys Asn Ser Asp Leu Val Ile Arg Ala Lys Phe Val Gly
35 40 45

Thr Pro Glu Val Asn Gln Thr Thr Leu Tyr Gln Arg Tyr Glu Ile Lys
50 55 60

Met Thr Lys Met Tyr Lys Gly Phe Gln Ala Leu Gly Asp Ala Ala Asp
65 70 75 80

Ile Arg Phe Val Tyr Thr Pro Ala Met Glu Ser Val Cys Gly Tyr Phe
85 90 95

His Arg Ser His Asn Arg Ser Glu Glu Phe Leu Ile Ala Gly Lys Leu
100 105 110

Gln Asp Gly Leu Leu His Ile Thr Thr Cys Ser Phe Val Ala Pro Trp
115 120 125

Asn Ser Leu Ser Leu Ala Gln Arg Arg Gly Phe Thr Lys Thr Tyr Thr
130 135 140

Val Gly Cys Glu Glu Cys Thr Val Phe Pro Cys Ser His Ser His Leu
145 150 155 160

Ser Ser Gly Gln

<210> 22
<211> 123
<212> PRT
<213> HUMAN

<400> 22
Met Ala Pro Phe Glu Pro Leu Ala Ser Gly Ile Leu Leu Leu Leu Trp
1 5 10 15

Leu Ile Ala Pro Ser Arg Ala Cys Thr Cys Val Pro Pro His Pro Gln
20 25 30

Thr Ala Phe Cys Asn Ser Asp Leu Val Ile Arg Ala Lys Phe Val Gly
35 40 45

Thr Pro Glu Val Asn Gln Thr Thr Leu Tyr Gln Arg Tyr Glu Ile Lys
50 55 60

Met Thr Lys Met Tyr Lys Gly Phe Gln Ala Leu Gly Asp Ala Ala Asp
65 70 75 80

Ile Arg Phe Val Tyr Thr Pro Ala Met Glu Ser Val Cys Gly Tyr Phe

```
                85                  90                  95

His Arg Ser His Asn Arg Ser Glu Glu Phe Leu Ile Ala Gly Lys Leu
            100                 105                 110

Gln Val Val Met Cys Lys Ser Pro Ser Val Val
        115                 120
```

<210> 23
<211> 211
<212> PRT
<213> HUMAN

<400> 23

```
Met Ala Pro Phe Glu Pro Leu Ala Ser Gly Ile Leu Leu Leu Leu Trp
  1               5                  10                  15

Leu Ile Ala Pro Ser Arg Ala Cys Thr Cys Val Pro Pro His Pro Gln
             20                  25                  30

Thr Ala Phe Cys Asn Ser Asp Leu Val Ile Arg Ala Lys Phe Val Gly
         35                  40                  45

Thr Pro Glu Val Asn Gln Thr Thr Leu Tyr Gln Arg Tyr Glu Ile Lys
     50                  55                  60

Met Thr Lys Met Tyr Lys Gly Phe Gln Ala Leu Gly Asp Ala Ala Asp
 65                  70                  75                  80

Ile Arg Phe Val Tyr Thr Pro Ala Met Glu Ser Val Cys Gly Tyr Phe
                85                  90                  95

His Arg Ser His Asn Arg Ser Glu Glu Phe Leu Ile Leu Leu Gly Lys
            100                 105                 110

Leu Gln Asp Gly Ile Phe Ala His Ser Leu Thr Cys Ser Phe Cys Trp
        115                 120                 125

Val Pro Trp Glu Asn Ser Leu Ser Leu Ala Gln Arg Arg Gly Phe Thr
    130                 135                 140

Lys Thr Tyr Thr Val Gly Cys Glu Glu Cys Thr Val Phe Pro Cys Leu
145                 150                 155                 160

Ser Ile Pro Cys Lys Leu Gln Ser Gly Thr His Cys Leu Trp Thr Asp
                165                 170                 175

Gln Leu Leu Gln Gly Ser Glu Lys Gly Phe Gln Ser Arg His Leu Ala
            180                 185                 190

Cys Leu Pro Arg Glu Pro Gly Leu Cys Thr Trp Gln Ser Leu Arg Ser
        195                 200                 205

Gln Ile Ala
    210
```

<210> 24

<211> 160
<212> PRT
<213> HUMAN

<400> 24
Met Ala Pro Phe Glu Pro Leu Ala Ser Gly Ile Leu Leu Leu Leu Trp
1 5 10 15

Leu Ile Ala Pro Ser Arg Ala Cys Thr Cys Val Pro Pro His Pro Gln
20 25 30

Thr Ala Phe Cys Asn Ser Asp Leu Val Ile Arg Ala Lys Phe Val Gly
35 40 45

Thr Pro Glu Val Asn Gln Thr Thr Leu Tyr Gln Arg Tyr Glu Ile Lys
50 55 60

Met Thr Lys Met Tyr Lys Gly Phe Gln Ala Leu Gly Asp Ala Ala Asp
65 70 75 80

Ile Arg Phe Val Tyr Thr Pro Ala Met Glu Ser Val Cys Gly Tyr Phe
85 90 95

His Arg Ser His Asn Arg Ser Glu Glu Phe Leu Ile Leu Ser Ile Pro
100 105 110

Cys Lys Leu Gln Ser Gly Thr His Cys Leu Trp Thr Asp Gln Leu Leu
115 120 125

Gln Gly Ser Glu Lys Gly Phe Gln Ser Arg His Leu Ala Cys Leu Pro
130 135 140

Arg Glu Pro Gly Leu Cys Thr Trp Gln Ser Leu Arg Ser Gln Ile Ala
145 150 155 160

<210> 25
<211> 197
<212> PRT
<213> HUMAN

<400> 25
Met Ala Pro Phe Glu Pro Leu Ala Ser Gly Ile Leu Leu Leu Leu Trp
1 5 10 15

Leu Ile Ala Pro Ser Arg Ala Cys Thr Cys Val Pro Pro His Pro Gln
20 25 30

Thr Ala Phe Cys Asn Ser Asp Leu Val Ile Arg Ala Lys Phe Val Gly
35 40 45

Thr Pro Glu Val Asn Gln Thr Thr Leu Tyr Gln Arg Tyr Glu Ile Lys
50 55 60

Met Thr Lys Met Tyr Lys Gly Phe Gln Ala Leu Gly Asp Ala Ala Asp
65 70 75 80

```
Ile Arg Phe Val Tyr Thr Pro Ala Met Glu Ser Val Cys Gly Tyr Phe
                  85                  90                  95

His Arg Ala Gly Lys Leu Gln Asp Gly Leu Leu His Ile Thr Thr Cys
             100                 105                 110

Ser Phe Val Ala Pro Trp Asn Ser Leu Ser Leu Ala Gln Arg Arg Gly
         115                 120                 125

Phe Thr Lys Thr Tyr Thr Val Gly Cys Glu Glu Cys Thr Val Phe Pro
     130                 135                 140

Cys Leu Ser Ile Pro Cys Lys Leu Gln Ser Gly Thr His Cys Leu Trp
145                 150                 155                 160

Thr Asp Gln Leu Leu Gln Gly Ser Glu Lys Gly Phe Gln Ser Arg His
                 165                 170                 175

Leu Ala Cys Leu Pro Arg Glu Pro Gly Leu Cys Thr Trp Gln Ser Leu
             180                 185                 190

Arg Ser Gln Ile Ala
         195
```

<210> 26
<211> 494
<212> PRT
<213> HUMAN

<400> 26

```
Met Arg Ala Leu Leu Ala Arg Leu Leu Leu Cys Val Leu Val Val Ser
  1               5                  10                  15

Asp Ser Lys Gly Ser Asn Glu Leu His Gln Val Pro Ser Asn Cys Asp
             20                  25                  30

Cys Leu Asn Gly Gly Thr Cys Val Ser Asn Lys Tyr Phe Ser Asn Ile
         35                  40                  45

His Trp Cys Asn Cys Pro Lys Lys Phe Gly Gly Gln His Cys Glu Ile
     50                  55                  60

Asp Lys Ser Lys Thr Cys Tyr Glu Gly Asn Gly His Phe Tyr Arg Gly
 65                  70                  75                  80

Lys Ala Ser Thr Asp Thr Met Gly Arg Pro Cys Leu Pro Trp Asn Ser
                 85                  90                  95

Ala Thr Val Leu Gln Gln Thr Tyr His Ala His Arg Ser Asp Ala Leu
             100                 105                 110

Gln Leu Gly Leu Gly Lys His Asn Tyr Cys Arg Glu Val Gly Ala Gln
         115                 120                 125

Gly Pro Lys Ala Leu Pro Thr Val Pro Arg Asn Leu Val Thr Ile Pro
     130                 135                 140

Phe Ser Gln Arg Ala Gly His Ser Thr Arg Glu Val Gln Pro Leu Val
145                 150                 155                 160
```

Glu Ser Ser Leu Arg Gly Gly Gly Arg Glu Gly Pro Leu Gly Trp Asn
165 170 175

Asp Ile Pro Tyr Leu Ser Val Leu Pro Gly Asn Pro Asp Asn Arg Arg
180 185 190

Arg Pro Trp Cys Tyr Val Gln Val Gly Leu Lys Pro Leu Val Gln Glu
195 200 205

Cys Met Val His Asp Cys Ala Asp Gly Lys Lys Pro Ser Ser Pro Pro
210 215 220

Glu Glu Leu Lys Phe Gln Cys Gly Gln Lys Thr Leu Arg Pro Arg Phe
225 230 235 240

Lys Ile Ile Gly Gly Glu Phe Thr Thr Ile Glu Asn Gln Pro Trp Phe
245 250 255

Ala Ala Ile Tyr Arg Arg His Arg Gly Gly Ser Val Thr Tyr Val Cys
260 265 270

Gly Gly Ser Leu Ile Ser Pro Cys Trp Val Ile Ser Ala Thr His Cys
275 280 285

Phe Ile Asp Tyr Pro Lys Lys Glu Asp Tyr Ile Val Tyr Leu Gly Arg
290 295 300

Ser Arg Leu Asn Ser Asn Thr Gln Gly Glu Met Lys Phe Glu Val Glu
305 310 315 320

Asn Leu Ile Leu His Lys Asp Tyr Ser Ala Asp Thr Leu Ala His His
325 330 335

Asn Asp Ile Ala Leu Leu Lys Ile Arg Ser Lys Glu Gly Arg Cys Ala
340 345 350

Gln Pro Ser Arg Thr Ile Gln Thr Ile Cys Leu Pro Ser Met Tyr Asn
355 360 365

Asp Pro Gln Phe Gly Thr Ser Cys Glu Ile Thr Gly Phe Gly Lys Glu
370 375 380

Asn Ser Thr Asp Tyr Leu Tyr Pro Glu Gln Leu Lys Met Thr Val Val
385 390 395 400

Lys Leu Ile Ser His Arg Glu Cys Gln Gln Pro His Tyr Tyr Gly Ser
405 410 415

Glu Val Thr Thr Lys Met Leu Cys Ala Ala Asp Pro Gln Trp Lys Thr
420 425 430

Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Ser Leu Gln
435 440 445

Gly Arg Met Thr Leu Thr Gly Ile Val Ser Trp Gly Arg Gly Cys Ala
450 455 460

Leu Lys Asp Lys Pro Gly Val Tyr Thr Arg Val Ser His Phe Leu Pro
465 470 475 480

```
Trp Ile Arg Ser His Thr Lys Glu Glu Asn Gly Leu Ala Leu
                485                 490
```

<210> 27
<211> 285
<212> PRT
<213> HUMAN

<400> 27

```
Met Gln Met Ser Pro Ala Leu Thr Cys Leu Val Leu Gly Leu Ala Leu
  1               5                  10                  15

Val Phe Gly Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala
             20                  25                  30

His Leu Ala Ser Asp Phe Gly Val Arg Val Phe Gln Gln Val Ala Gln
         35                  40                  45

Ala Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser
     50                  55                  60

Val Leu Ala Met Leu Gln Leu Thr Thr Gly Gly Glu Thr Gln Gln Gln
 65                  70                  75                  80

Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro
                85                  90                  95

Ala Leu Arg His Leu Tyr Lys Glu Leu Met Gly Pro Trp Asn Lys Asp
            100                 105                 110

Glu Ile Ser Thr Thr Asp Ala Ile Phe Val Gln Arg Asp Leu Lys Leu
        115                 120                 125

Val Gln Gly Phe Met Pro His Phe Phe Arg Leu Phe Arg Ser Thr Val
    130                 135                 140

Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe Ile Ile Asn
145                 150                 155                 160

Asp Trp Val Lys Thr His Thr Lys Gly Met Ile Ser Asn Leu Leu Gly
                165                 170                 175

Lys Gly Ala Val Asp Gln Leu Thr Arg Leu Val Leu Val Asn Ala Leu
            180                 185                 190

Tyr Phe Asn Gly Gln Trp Lys Thr Pro Phe Pro Asp Ser Ser Thr His
        195                 200                 205

Arg Arg Leu Phe His Lys Ser Asp Gly Ser Thr Val Ser Val Pro Met
    210                 215                 220

Met Ala Gln Thr Asn Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp
225                 230                 235                 240

Gly His Tyr Tyr Asp Ile Leu Glu Leu Pro Tyr His Gly Asp Thr Leu
                245                 250                 255

Ser Met Phe Ile Ala Ala Asp Leu Val Pro Thr Glu Ala Leu Cys Arg
            260                 265                 270
```

```
Met Glu Leu Arg Gly Leu Gln Glu Leu Leu Cys Ala Trp
        275                 280                 285
```

<210> 28
<211> 399
<212> PRT
<213> HUMAN

<400> 28

```
Met Gln Met Ser Pro Ala Leu Thr Cys Leu Val Leu Gly Leu Ala Leu
  1               5                  10                  15

Val Phe Gly Glu Gly Ser Ala Val His His Pro Pro Ser Tyr Val Ala
             20                  25                  30

His Leu Ala Ser Asp Phe Gly Val Arg Val Phe Gln Gln Val Ala Gln
         35                  40                  45

Ala Ser Lys Asp Arg Asn Val Val Phe Ser Pro Tyr Gly Val Ala Ser
     50                  55                  60

Val Leu Ala Met Leu Gln Leu Thr Thr Gly Gly Glu Thr Gln Gln Gln
 65                  70                  75                  80

Ile Gln Ala Ala Met Gly Phe Lys Ile Asp Asp Lys Gly Met Ala Pro
                 85                  90                  95

Ala Leu Arg His Leu Tyr Lys Glu Leu Met Gly Pro Trp Asn Lys Asp
            100                 105                 110

Glu Ile Ser Thr Thr Asp Ala Ile Phe Val Gln Arg Asp Leu Lys Leu
        115                 120                 125

Val Gln Gly Phe Met Pro His Phe Phe Arg Leu Phe Arg Ser Thr Val
    130                 135                 140

Lys Gln Val Asp Phe Ser Glu Val Glu Arg Ala Arg Phe Ile Ile Asn
145                 150                 155                 160

Asp Trp Val Lys Thr His Thr Lys Gly Met Ile Ser Asn Leu Leu Gly
                165                 170                 175

Lys Gly Ala Val Asp Gln Leu Thr Arg Leu Val Leu Val Asn Ala Leu
            180                 185                 190

Tyr Phe Asn Gly Gln Trp Lys Thr Pro Phe Pro Asp Ser Ser Thr His
        195                 200                 205

Arg Arg Leu Phe His Lys Ser Asp Gly Ser Thr Val Ser Val Pro Met
    210                 215                 220

Met Ala Gln Thr Asn Lys Phe Asn Tyr Thr Glu Phe Thr Thr Pro Asp
225                 230                 235                 240

Gly His Tyr Tyr Asp Ile Leu Glu Leu Pro Tyr His Gly Asp Thr Leu
                245                 250                 255

Ser Met Phe Ile Ala Ala Pro Tyr Glu Lys Glu Val Pro Leu Ser Ala
```

```
            260                 265                 270

Leu Thr Asn Ile Leu Ser Ala Gln Leu Ile Ser His Trp Lys Gly Asn
        275                 280                 285

Met Thr Arg Leu Pro Arg Leu Leu Val Leu Pro Lys Phe Ser Leu Glu
    290                 295                 300

Thr Glu Val Asp Leu Arg Lys Pro Leu Glu Asn Leu Gly Met Thr Asp
305                 310                 315                 320

Met Phe Arg Gln Phe Gln Ala Asp Phe Thr Ser Leu Ser Asp Gln Glu
                325                 330                 335

Pro Leu His Val Ala Gln Ala Leu Gln Lys Val Lys Ile Glu Val Asn
            340                 345                 350

Glu Ser Gly Thr Val Ala Ser Ser Ser Thr Ala Val Ile Val Ser Ala
        355                 360                 365

Arg Met Ala Pro Glu Glu Ile Ile Met Asp Arg Pro Phe Leu Phe Val
    370                 375                 380

Val Pro Pro Gln Lys Gln Cys Ala Trp Val Ile Leu Glu Cys Arg
385                 390                 395
```

<210> 29
<211> 317
<212> PRT
<213> HUMAN

<400> 29

```
Met Thr Ala Ala Ser Met Gly Pro Val Arg Val Ala Phe Val Val Leu
  1               5                  10                  15

Leu Ala Leu Cys Ser Arg Pro Ala Val Gly Gln Asn Cys Ser Gly Pro
            20                  25                  30

Cys Arg Cys Pro Asp Glu Pro Ala Pro Arg Cys Pro Ala Gly Val Ser
        35                  40                  45

Leu Val Leu Asp Gly Cys Gly Cys Cys Arg Val Cys Ala Lys Gln Leu
    50                  55                  60

Gly Glu Leu Cys Thr Glu Arg Asp Pro Cys Asp Pro His Lys Gly Leu
65                  70                  75                  80

Phe Cys Asp Phe Gly Ser Pro Ala Asn Arg Lys Ile Gly Val Cys Thr
                85                  90                  95

Ala Lys Asp Gly Ala Pro Cys Ile Phe Gly Gly Thr Val Tyr Arg Ser
            100                 105                 110

Gly Glu Ser Phe Gln Ser Ser Cys Lys Tyr Gln Cys Thr Cys Leu Asp
        115                 120                 125

Gly Ala Val Gly Cys Met Pro Leu Cys Ser Met Asp Val Arg Leu Pro
    130                 135                 140
```

```
Ser Pro Asp Cys Pro Leu Pro Leu Glu Asp Thr Phe Gly Pro Asp Pro
145                 150                 155                 160

Thr Met Ile Arg Ala Asn Cys Leu Val Gln Thr Thr Glu Trp Ser Ala
                165                 170                 175

Cys Ser Lys Thr Cys Gly Met Gly Ile Ser Thr Arg Val Thr Asn Asp
            180                 185                 190

Asn Ala Ser Cys Arg Leu Glu Lys Gln Ser Arg Leu Cys Met Val Arg
        195                 200                 205

Pro Cys Glu Ser Asp Leu Glu Glu Asn Ile Lys Lys Gly Lys Lys Cys
    210                 215                 220

Ile Arg Thr Pro Lys Ile Ser Lys Pro Ile Lys Phe Glu Leu Ser Gly
225                 230                 235                 240

Cys Thr Ser Met Lys Thr Tyr Arg Ala Lys Phe Cys Gly Val Cys Thr
                245                 250                 255

Asp Gly Arg Cys Cys Thr Pro His Arg Thr Thr Thr Leu Pro Val Glu
            260                 265                 270

Phe Lys Cys Pro Asp Gly Glu Val Met Lys Lys Asn Met Met Phe Ile
        275                 280                 285

Lys Thr Cys Ala Cys His Tyr Asn Cys Pro Gly Asp Asn Asp Ile Phe
    290                 295                 300

Glu Ser Leu Tyr Tyr Arg Lys Met Tyr Gly Asp Met Ala
305                 310                 315
```

<210> 30
<211> 342
<212> PRT
<213> HUMAN

<400> 30

```
Asn Met Glu Asn Ser Leu Arg Cys Val Trp Val Pro Lys Leu Ala Phe
  1               5                  10                  15

Val Leu Phe Gly Ala Ser Leu Leu Ser Ala His Leu Gln Val Thr Gly
             20                  25                  30

Phe Gln Ile Lys Ala Phe Thr Ala Leu Arg Phe Leu Ser Glu Pro Ser
         35                  40                  45

Asp Ala Val Thr Met Arg Gly Gly Asn Val Leu Leu Asp Cys Ser Ala
     50                  55                  60

Glu Ser Asp Arg Gly Val Pro Val Ile Lys Trp Lys Lys Asp Ala Ile
 65                  70                  75                  80

His Leu Ala Leu Gly Met Asp Glu Arg Lys Gln Gln Leu Ser Asn Gly
                 85                  90                  95

Ser Leu Leu Ile Gln Asn Ile Leu His Ser Arg His His Lys Pro Asp
            100                 105                 110
```

```
Glu Gly Leu Tyr Gln Cys Glu Ala Ser Leu Gly Asp Ser Gly Ser Ile
        115                 120                 125

Ile Ser Arg Thr Ala Lys Val Ala Val Ala Gly Pro Leu Arg Phe Leu
    130                 135                 140

Ser Gln Thr Glu Ser Val Thr Ala Phe Met Gly Asp Thr Val Leu Leu
145                 150                 155                 160

Lys Cys Glu Val Ile Gly Glu Pro Met Pro Thr Ile His Trp Gln Lys
            165                 170                 175

Asn Gln Gln Asp Leu Thr Pro Ile Pro Gly Asp Ser Arg Val Val Val
        180                 185                 190

Leu Pro Ser Gly Ala Leu Gln Ile Ser Arg Leu Gln Pro Gly Asp Ile
        195                 200                 205

Gly Ile Tyr Arg Cys Ser Ala Arg Asn Pro Ala Ser Ser Arg Thr Gly
    210                 215                 220

Asn Glu Ala Glu Val Arg Ile Leu Ser Asp Pro Gly Leu His Arg Gln
225                 230                 235                 240

Leu Tyr Phe Leu Gln Arg Pro Ser Asn Val Val Ala Ile Glu Gly Lys
                245                 250                 255

Asp Ala Val Leu Glu Cys Cys Val Ser Gly Tyr Pro Pro Pro Ser Phe
        260                 265                 270

Thr Trp Leu Arg Gly Glu Glu Val Ile Gln Leu Arg Ser Lys Lys Tyr
        275                 280                 285

Ser Leu Leu Gly Gly Ser Asn Leu Leu Ile Ser Asn Val Thr Asp Asp
    290                 295                 300

Asp Ser Gly Met Tyr Thr Cys Val Val Thr Tyr Lys Asn Glu Asn Ile
305                 310                 315                 320

Ser Ala Ser Ala Glu Leu Thr Val Leu Val Ile Ile Asp Lys Val Leu
                325                 330                 335

Val Asp Thr Phe Trp Val
        340
```

<210> 31
<211> 1433
<212> PRT
<213> HUMAN

<400> 31

```
Asn Met Glu Asn Ser Leu Arg Cys Val Trp Val Pro Lys Leu Ala Phe
  1               5                  10                  15

Val Leu Phe Gly Ala Ser Leu Leu Ser Ala His Leu Gln Val Thr Gly
            20                  25                  30

Phe Gln Ile Lys Ala Phe Thr Ala Leu Arg Phe Leu Ser Glu Pro Ser
```

```
          35                  40                  45

Asp Ala Val Thr Met Arg Gly Gly Asn Val Leu Leu Asp Cys Ser Ala
     50                  55                  60

Glu Ser Asp Arg Gly Val Pro Val Ile Lys Trp Lys Lys Asp Ala Ile
 65                  70                  75                  80

His Leu Ala Leu Gly Met Asp Glu Arg Lys Gln Gln Leu Ser Asn Gly
                 85                  90                  95

Ser Leu Leu Ile Gln Asn Ile Leu His Ser Arg His His Lys Pro Asp
            100                 105                 110

Glu Gly Leu Tyr Gln Cys Glu Ala Ser Leu Gly Asp Ser Gly Ser Ile
        115                 120                 125

Ile Ser Arg Thr Ala Lys Val Ala Val Ala Gly Pro Leu Arg Phe Leu
    130                 135                 140

Ser Gln Thr Glu Ser Val Thr Ala Phe Met Gly Asp Thr Val Leu Leu
145                 150                 155                 160

Lys Cys Glu Val Ile Gly Glu Pro Met Pro Thr Ile His Trp Gln Lys
                165                 170                 175

Asn Gln Gln Asp Leu Thr Pro Ile Pro Gly Asp Ser Arg Val Val Val
            180                 185                 190

Leu Pro Ser Gly Ala Leu Gln Ile Ser Arg Leu Gln Pro Gly Asp Ile
        195                 200                 205

Gly Ile Tyr Arg Cys Ser Ala Arg Asn Pro Ala Ser Ser Arg Thr Gly
    210                 215                 220

Asn Glu Ala Glu Val Arg Ile Leu Ser Asp Pro Gly Leu His Arg Gln
225                 230                 235                 240

Leu Tyr Phe Leu Gln Arg Pro Ser Asn Val Val Ala Ile Glu Gly Lys
                245                 250                 255

Asp Ala Val Leu Glu Cys Cys Val Ser Gly Tyr Pro Pro Pro Ser Phe
            260                 265                 270

Thr Trp Leu Arg Gly Glu Glu Val Ile Gln Leu Arg Ser Lys Lys Tyr
        275                 280                 285

Ser Leu Leu Gly Gly Ser Asn Leu Leu Ile Ser Asn Val Thr Asp Asp
    290                 295                 300

Asp Ser Gly Met Tyr Thr Cys Val Val Thr Tyr Lys Asn Glu Asn Ile
305                 310                 315                 320

Ser Ala Ser Ala Glu Leu Thr Val Leu Val Pro Pro Trp Phe Leu Asn
                325                 330                 335

His Pro Ser Asn Leu Tyr Ala Tyr Glu Ser Met Asp Ile Glu Phe Glu
            340                 345                 350

Cys Thr Val Ser Gly Lys Pro Val Pro Thr Val Asn Trp Met Lys Asn
        355                 360                 365
```

```
Gly Asp Val Val Ile Pro Ser Asp Tyr Phe Gln Ile Val Gly Gly Ser
    370                 375                 380

Asn Leu Arg Ile Leu Gly Val Val Lys Ser Asp Glu Gly Phe Tyr Gln
385                 390                 395                 400

Cys Val Ala Glu Asn Glu Ala Gly Asn Ala Gln Thr Ser Ala Gln Leu
                405                 410                 415

Ile Val Pro Lys Pro Ala Ile Pro Ser Ser Ser Val Leu Pro Ser Ala
            420                 425                 430

Pro Arg Asp Val Val Pro Val Leu Val Ser Ser Arg Phe Val Arg Leu
        435                 440                 445

Ser Trp Arg Pro Pro Ala Glu Ala Lys Gly Asn Ile Gln Thr Phe Thr
    450                 455                 460

Val Phe Phe Ser Arg Glu Gly Asp Asn Arg Glu Arg Ala Leu Asn Thr
465                 470                 475                 480

Thr Gln Pro Gly Ser Leu Gln Leu Thr Val Gly Asn Leu Lys Pro Glu
                485                 490                 495

Ala Met Tyr Thr Phe Arg Val Val Ala Tyr Asn Glu Trp Gly Pro Gly
            500                 505                 510

Glu Ser Ser Gln Pro Ile Lys Val Ala Thr Gln Pro Glu Leu Gln Val
        515                 520                 525

Pro Gly Pro Val Glu Asn Leu Gln Ala Val Ser Thr Ser Pro Thr Ser
    530                 535                 540

Ile Leu Ile Thr Trp Glu Pro Pro Ala Tyr Ala Asn Gly Pro Val Gln
545                 550                 555                 560

Gly Tyr Arg Leu Phe Cys Thr Glu Val Ser Thr Gly Lys Glu Gln Asn
                565                 570                 575

Ile Glu Val Asp Gly Leu Ser Tyr Lys Leu Glu Gly Leu Lys Lys Phe
            580                 585                 590

Thr Glu Tyr Ser Leu Arg Phe Leu Ala Tyr Asn Arg Tyr Gly Pro Gly
        595                 600                 605

Val Ser Thr Asp Asp Ile Thr Val Val Thr Leu Ser Asp Val Pro Ser
    610                 615                 620

Ala Pro Pro Gln Asn Val Ser Leu Glu Val Val Asn Ser Arg Ser Ile
625                 630                 635                 640

Lys Val Ser Trp Leu Pro Pro Pro Ser Gly Thr Gln Asn Gly Phe Ile
                645                 650                 655

Thr Gly Tyr Lys Ile Arg His Arg Lys Thr Thr Arg Arg Gly Glu Met
            660                 665                 670

Glu Thr Leu Glu Pro Asn Asn Leu Trp Tyr Leu Phe Thr Gly Leu Glu
        675                 680                 685
```

Lys Gly Ser Gln Tyr Ser Phe Gln Val Ser Ala Met Thr Val Asn Gly
690 695 700

Thr Gly Pro Pro Ser Asn Trp Tyr Thr Ala Glu Thr Pro Glu Asn Asp
705 710 715 720

Leu Asp Glu Ser Gln Val Pro Asp Gln Pro Ser Ser Leu His Val Arg
725 730 735

Pro Gln Thr Asn Cys Ile Ile Met Ser Trp Thr Pro Pro Leu Asn Pro
740 745 750

Asn Ile Val Val Arg Gly Tyr Ile Ile Gly Tyr Gly Val Gly Ser Pro
755 760 765

Tyr Ala Glu Thr Val Arg Val Asp Ser Lys Gln Arg Tyr Tyr Ser Ile
770 775 780

Glu Arg Leu Glu Ser Ser Ser His Tyr Val Ile Ser Leu Lys Ala Phe
785 790 795 800

Asn Asn Ala Gly Glu Gly Val Pro Leu Tyr Glu Ser Ala Thr Thr Arg
805 810 815

Ser Ile Thr Asp Pro Thr Asp Pro Val Asp Tyr Tyr Pro Leu Leu Asp
820 825 830

Asp Phe Pro Thr Ser Val Pro Asp Leu Ser Thr Pro Met Leu Pro Pro
835 840 845

Val Gly Val Gln Ala Val Ala Leu Thr His Asp Ala Val Arg Val Ser
850 855 860

Trp Ala Asp Asn Ser Val Pro Lys Asn Gln Lys Thr Ser Glu Val Arg
865 870 875 880

Leu Tyr Thr Val Arg Trp Arg Thr Ser Phe Ser Ala Ser Ala Lys Tyr
885 890 895

Lys Ser Glu Asp Thr Thr Ser Leu Ser Tyr Thr Ala Thr Gly Leu Lys
900 905 910

Pro Asn Thr Met Tyr Glu Phe Ser Val Met Val Thr Lys Asn Arg Arg
915 920 925

Ser Ser Thr Trp Ser Met Thr Ala His Ala Thr Thr Tyr Glu Ala Ala
930 935 940

Pro Thr Ser Ala Pro Lys Asp Phe Thr Val Ile Thr Arg Glu Gly Lys
945 950 955 960

Pro Arg Ala Val Ile Val Ser Trp Gln Pro Pro Leu Glu Ala Asn Gly
965 970 975

Lys Ile Thr Ala Tyr Ile Leu Phe Tyr Thr Leu Asp Lys Asn Ile Pro
980 985 990

Ile Asp Asp Trp Ile Met Glu Thr Ile Ser Gly Asp Arg Leu Thr His
995 1000 1005

Gln Ile Met Asp Leu Asn Leu Asp Thr Met Tyr Tyr Phe Arg Ile Gln

```
1010                1015                1020

Ala Arg Asn Ser Lys Gly Val Gly Pro Leu Ser Asp Pro Ile Leu Phe
1025                1030                1035                1040

Arg Thr Leu Lys Val Glu His Pro Asp Lys Met Ala Asn Asp Gln Gly
                1045                1050                1055

Arg His Gly Asp Gly Gly Tyr Trp Pro Val Asp Thr Asn Leu Ile Asp
            1060                1065                1070

Arg Ser Thr Leu Asn Glu Pro Pro Ile Gly Gln Met His Pro Pro His
        1075                1080                1085

Gly Ser Val Thr Pro Gln Lys Asn Ser Asn Leu Leu Val Ile Ile Val
    1090                1095                1100

Val Thr Val Gly Val Ile Thr Val Leu Val Val Val Ile Val Ala Val
1105                1110                1115                1120

Ile Cys Thr Arg Arg Ser Ser Ala Gln Gln Arg Lys Lys Arg Ala Thr
                1125                1130                1135

His Ser Ala Gly Lys Arg Lys Gly Ser Gln Lys Asp Leu Arg Pro Pro
            1140                1145                1150

Asp Leu Trp Ile His His Glu Glu Met Glu Met Lys Asn Ile Glu Lys
        1155                1160                1165

Pro Ser Gly Thr Asp Pro Ala Gly Arg Asp Ser Pro Ile Gln Ser Cys
    1170                1175                1180

Gln Asp Leu Thr Pro Val Ser His Ser Gln Ser Glu Thr Gln Leu Gly
1185                1190                1195                1200

Ser Lys Ser Thr Ser His Ser Gly Gln Asp Thr Glu Glu Ala Gly Ser
                1205                1210                1215

Ser Met Ser Thr Leu Glu Arg Ser Leu Ala Ala Arg Arg Ala Pro Arg
            1220                1225                1230

Ala Lys Leu Met Ile Pro Met Asp Ala Gln Ser Asn Asn Pro Ala Val
        1235                1240                1245

Val Ser Ala Ile Pro Val Pro Thr Leu Glu Ser Ala Gln Tyr Pro Gly
    1250                1255                1260

Ile Leu Pro Ser Pro Thr Cys Gly Tyr Pro His Pro Gln Phe Thr Leu
1265                1270                1275                1280

Arg Pro Val Pro Phe Pro Thr Leu Ser Val Asp Arg Gly Phe Gly Ala
                1285                1290                1295

Gly Arg Ser Gln Ser Val Ser Glu Gly Pro Thr Thr Gln Gln Pro Pro
            1300                1305                1310

Met Leu Pro Pro Ser Gln Pro Glu His Ser Ser Ser Glu Glu Ala Pro
        1315                1320                1325

Ser Arg Thr Ile Pro Thr Ala Cys Val Arg Pro Thr His Pro Leu Arg
    1330                1335                1340
```

```
Ser Phe Ala Asn Pro Leu Leu Pro Pro Pro Met Ser Ala Ile Glu Pro
1345                1350                1355                1360

Lys Val Pro Tyr Thr Pro Leu Leu Ser Gln Pro Gly Pro Thr Leu Pro
                1365                1370                1375

Lys Thr His Val Lys Thr Ala Ser Leu Gly Leu Ala Gly Lys Ala Arg
            1380                1385                1390

Ser Pro Leu Leu Pro Val Ser Val Pro Thr Ala Pro Glu Val Ser Glu
        1395                1400                1405

Glu Ser His Lys Pro Thr Glu Asp Ser Ala Asn Val Ser Ala Ser Leu
    1410                1415                1420

Lys Phe Met Leu His Gln Gly Thr Asp
1425                1430
```

<210> 32
<211> 865
<212> PRT
<213> HUMAN

<400> 32

```
Met Pro Gly Lys Arg Gly Leu Gly Trp Trp Trp Ala Arg Leu Pro Leu
  1               5                  10                  15

Cys Leu Leu Leu Ser Leu Tyr Gly Pro Trp Met Pro Ser Ser Leu Gly
             20                  25                  30

Lys Pro Lys Gly His Pro His Met Asn Ser Ile Arg Ile Asp Gly Asp
         35                  40                  45

Ile Thr Leu Gly Gly Leu Phe Pro Val His Gly Arg Gly Ser Glu Gly
     50                  55                  60

Lys Pro Cys Gly Glu Leu Lys Lys Glu Lys Gly Ile His Arg Leu Glu
 65                  70                  75                  80

Ala Met Leu Phe Ala Leu Asp Arg Ile Asn Asn Asp Pro Asp Leu Leu
                 85                  90                  95

Pro Asn Ile Thr Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp
             100                 105                 110

Thr His Ala Leu Glu Gln Ser Leu Thr Phe Val Gln Ala Leu Ile Glu
         115                 120                 125

Lys Asp Gly Thr Glu Val Arg Cys Gly Ser Gly Gly Pro Pro Ile Ile
     130                 135                 140

Thr Lys Pro Glu Arg Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser
145                 150                 155                 160

Val Ser Ile Met Val Ala Asn Ile Leu Arg Leu Phe Lys Ile Pro Gln
                165                 170                 175

Ile Ser Tyr Ala Ser Thr Ala Pro Asp Leu Ser Asp Asn Ser Arg Tyr
```

```
            180                 185                 190

Asp Phe Phe Ser Arg Val Val Pro Ser Asp Thr Tyr Gln Ala Gln Ala
        195                 200                 205

Met Val Asp Ile Val Arg Ala Leu Lys Trp Asn Tyr Val Ser Thr Val
    210                 215                 220

Ala Ser Glu Gly Ser Tyr Gly Glu Ser Gly Val Glu Ala Phe Ile Gln
225                 230                 235                 240

Lys Ser Arg Glu Asp Gly Gly Val Cys Ile Ala Gln Ser Val Lys Ile
                245                 250                 255

Pro Arg Glu Pro Lys Ala Gly Glu Phe Asp Lys Ile Ile Arg Arg Leu
            260                 265                 270

Leu Glu Thr Ser Asn Ala Arg Ala Val Ile Ile Phe Ala Asn Glu Asp
        275                 280                 285

Asp Ile Arg Arg Val Leu Glu Ala Ala Arg Arg Ala Asn Gln Thr Gly
    290                 295                 300

His Phe Phe Trp Met Gly Ser Asp Ser Trp Gly Ser Lys Ile Ala Pro
305                 310                 315                 320

Val Leu His Leu Glu Glu Val Ala Glu Gly Ala Val Thr Ile Leu Pro
                325                 330                 335

Lys Arg Met Ser Val Arg Asp Arg Glu Arg Ile Gly Gln Asp Ser Ala
            340                 345                 350

Tyr Glu Gln Glu Gly Lys Val Gln Phe Val Ile Asp Ala Val Tyr Ala
        355                 360                 365

Met Gly His Ala Leu His Ala Met His Arg Asp Leu Cys Pro Gly Arg
    370                 375                 380

Val Gly Leu Cys Pro Arg Met Asp Pro Val Asp Gly Thr Gln Leu Leu
385                 390                 395                 400

Lys Tyr Ile Arg Asn Val Asn Phe Ser Gly Ile Ala Gly Asn Pro Val
                405                 410                 415

Thr Phe Asn Glu Asn Gly Asp Ala Pro Gly Arg Tyr Asp Ile Tyr Gln
            420                 425                 430

Tyr Gln Leu Arg Asn Asp Ser Ala Glu Tyr Lys Val Ile Gly Ser Trp
        435                 440                 445

Thr Asp His Leu His Leu Arg Ile Glu Arg Met His Trp Pro Gly Ser
    450                 455                 460

Gly Gln Gln Leu Pro Arg Ser Ile Cys Ser Leu Pro Cys Gln Pro Gly
465                 470                 475                 480

Glu Arg Lys Lys Thr Val Lys Gly Met Pro Cys Cys Trp His Cys Glu
                485                 490                 495

Pro Cys Thr Gly Tyr Gln Tyr Gln Val Asp Arg Tyr Thr Cys Lys Thr
            500                 505                 510
```

Cys Pro Tyr Asp Met Arg Pro Thr Glu Asn Arg Thr Gly Cys Arg Pro
515 520 525

Ile Pro Ile Ile Lys Leu Glu Trp Gly Ser Pro Trp Ala Val Leu Pro
530 535 540

Leu Phe Leu Ala Val Val Gly Ile Ala Ala Thr Leu Phe Val Val Ile
545 550 555 560

Thr Phe Val Arg Tyr Asn Asp Thr Pro Ile Val Lys Ala Ser Gly Arg
565 570 575

Glu Leu Ser Tyr Val Leu Leu Ala Gly Ile Phe Leu Cys Tyr Ala Thr
580 585 590

Thr Phe Leu Met Ile Ala Glu Pro Asp Leu Gly Thr Cys Ser Leu Arg
595 600 605

Arg Ile Phe Leu Gly Leu Gly Met Ser Ile Ser Tyr Ala Ala Leu Leu
610 615 620

Thr Lys Thr Asn Arg Ile Tyr Arg Ile Phe Glu Gln Gly Lys Arg Ser
625 630 635 640

Val Ser Ala Pro Arg Phe Ile Ser Pro Ala Ser Gln Leu Ala Ile Thr
645 650 655

Phe Ser Leu Ile Ser Leu Gln Leu Leu Gly Ile Cys Val Trp Phe Val
660 665 670

Val Asp Pro Ser His Ser Val Val Asp Phe Gln Asp Gln Arg Thr Leu
675 680 685

Asp Pro Arg Phe Ala Arg Gly Val Leu Lys Cys Asp Ile Ser Asp Leu
690 695 700

Ser Leu Ile Cys Leu Leu Gly Tyr Ser Met Leu Leu Met Val Thr Cys
705 710 715 720

Thr Val Tyr Ala Ile Lys Thr Arg Gly Val Pro Glu Thr Phe Asn Glu
725 730 735

Ala Lys Pro Ile Gly Phe Thr Met Tyr Thr Thr Cys Ile Val Trp Leu
740 745 750

Ala Phe Ile Pro Ile Phe Phe Gly Thr Ser Gln Ser Ala Asp Lys Leu
755 760 765

Tyr Ile Gln Thr Thr Thr Leu Thr Val Ser Val Ser Leu Ser Ala Ser
770 775 780

Val Ser Leu Gly Met Leu Tyr Met Pro Lys Val Tyr Ile Ile Leu Phe
785 790 795 800

His Pro Glu Gln Asn Val Pro Lys Arg Lys Arg Ser Leu Lys Ala Val
805 810 815

Val Thr Ala Ala Thr Met Ser Asn Lys Phe Thr Gln Lys Gly Asn Phe
820 825 830

```
Arg Pro Asn Gly Glu Ala Lys Ser Glu Leu Cys Glu Asn Leu Glu Ala
        835                 840                 845

Pro Ala Leu Ala Thr Lys Gln Thr Tyr Val Thr Tyr Thr Asn His Ala
    850                 855                 860

Ile
865
```

<210> 33
<211> 845
<212> PRT
<213> HUMAN

<400> 33

```
Met Glu Thr Lys Gly Tyr His Ser Leu Pro Glu Gly Leu Asp Met Glu
  1               5                  10                  15

Arg Arg Trp Gly Gln Val Ser Gln Ala Val Glu Arg Ser Ser Leu Gly
             20                  25                  30

Pro Thr Glu Arg Thr Asp Glu Asn Asn Tyr Met Glu Ile Val Asn Val
         35                  40                  45

Ser Cys Val Ser Gly Ala Ile Pro Asn Asn Ser Thr Gln Gly Ser Ser
     50                  55                  60

Lys Glu Lys Gln Glu Leu Leu Pro Cys Leu Gln Gln Asp Asn Asn Arg
 65                  70                  75                  80

Pro Gly Ile Leu Thr Ser Asp Ile Lys Thr Glu Leu Glu Ser Lys Glu
                 85                  90                  95

Leu Ser Ala Thr Val Ala Glu Ser Met Gly Leu Tyr Met Asp Ser Val
            100                 105                 110

Arg Asp Ala Asp Tyr Ser Tyr Glu Gln Gln Asn Gln Gln Gly Ser Met
        115                 120                 125

Ser Pro Ala Lys Ile Tyr Gln Asn Val Glu Gln Leu Val Lys Phe Tyr
    130                 135                 140

Lys Gly Asn Gly His Arg Pro Ser Thr Leu Ser Cys Val Asn Thr Pro
145                 150                 155                 160

Leu Arg Ser Phe Met Ser Asp Ser Gly Ser Ser Val Asn Gly Gly Val
                165                 170                 175

Met Arg Ala Ile Val Lys Ser Pro Ile Met Cys His Glu Lys Ser Pro
            180                 185                 190

Ser Val Cys Ser Pro Leu Asn Met Thr Ser Ser Val Cys Ser Pro Ala
        195                 200                 205

Gly Ile Asn Ser Val Ser Ser Thr Thr Ala Ser Phe Gly Ser Phe Pro
    210                 215                 220

Val His Ser Pro Ile Thr Gln Gly Thr Pro Leu Thr Cys Ser Pro Asn
225                 230                 235                 240
```

Ala Glu Asn Arg Gly Ser Arg Ser His Ser Pro Ala His Ala Ser Asn
245 250 255

Val Gly Ser Pro Leu Ser Ser Pro Leu Ser Ser Met Lys Ser Ser Ile
260 265 270

Ser Ser Pro Pro Ser His Cys Ser Val Lys Ser Pro Val Ser Ser Pro
275 280 285

Asn Asn Val Thr Leu Arg Ser Ser Val Ser Ser Pro Ala Asn Ile Asn
290 295 300

Asn Ser Arg Cys Ser Val Ser Ser Pro Ser Asn Thr Asn Asn Arg Ser
305 310 315 320

Thr Leu Ser Ser Pro Ala Ala Ser Thr Val Gly Ser Ile Cys Ser Pro
325 330 335

Val Asn Asn Ala Phe Ser Tyr Thr Ala Ser Gly Thr Ser Ala Gly Ser
340 345 350

Ser Thr Leu Arg Asp Val Val Pro Ser Pro Asp Thr Gln Glu Lys Gly
355 360 365

Ala Gln Glu Val Pro Phe Pro Lys Thr Glu Glu Val Glu Ser Ala Ile
370 375 380

Ser Asn Gly Val Thr Gly Gln Leu Asn Ile Val Gln Tyr Ile Lys Pro
385 390 395 400

Glu Pro Asp Gly Ala Phe Ser Ser Ser Cys Leu Gly Gly Asn Ser Lys
405 410 415

Ile Asn Ser Asp Ser Ser Phe Ser Val Pro Ile Lys Gln Glu Ser Thr
420 425 430

Lys His Ser Cys Ser Gly Thr Ser Phe Lys Gly Asn Pro Thr Val Asn
435 440 445

Pro Phe Pro Phe Met Asp Gly Ser Tyr Phe Ser Phe Met Asp Asp Lys
450 455 460

Asp Tyr Tyr Ser Leu Ser Gly Ile Leu Gly Pro Pro Val Pro Gly Phe
465 470 475 480

Asp Gly Asn Cys Glu Gly Ser Gly Phe Pro Val Gly Ile Lys Gln Glu
485 490 495

Pro Asp Asp Gly Ser Tyr Tyr Pro Glu Ala Ser Ile Pro Ser Ser Ala
500 505 510

Ile Val Gly Val Asn Ser Gly Gly Gln Ser Phe His Tyr Arg Ile Gly
515 520 525

Ala Gln Gly Thr Ile Ser Leu Ser Arg Ser Ala Arg Asp Gln Ser Phe
530 535 540

Gln His Leu Ser Ser Phe Pro Pro Val Asn Thr Leu Val Glu Ser Trp
545 550 555 560

```
Lys Ser His Gly Asp Leu Ser Ser Arg Arg Ser Asp Gly Tyr Pro Val
                565                 570                 575

Leu Glu Tyr Ile Pro Glu Asn Val Ser Ser Ser Thr Leu Arg Ser Val
            580                 585                 590

Ser Thr Gly Ser Ser Arg Pro Ser Lys Ile Cys Leu Val Cys Gly Asp
        595                 600                 605

Glu Ala Ser Gly Cys His Tyr Gly Val Val Thr Cys Gly Ser Cys Lys
    610                 615                 620

Val Phe Phe Lys Arg Ala Val Glu Gly Gln His Asn Tyr Leu Cys Ala
625                 630                 635                 640

Gly Arg Asn Asp Cys Ile Ile Asp Lys Ile Arg Arg Lys Asn Cys Pro
                645                 650                 655

Ala Cys Arg Leu Gln Lys Cys Leu Gln Ala Gly Met Asn Leu Gly Ala
            660                 665                 670

Arg Lys Ser Lys Lys Leu Gly Lys Leu Lys Gly Ile His Glu Glu Gln
        675                 680                 685

Pro Gln Gln Gln Gln Pro Pro Pro Pro Pro Pro Pro Pro Gln Ser Pro
    690                 695                 700

Glu Glu Gly Thr Thr Tyr Ile Ala Pro Ala Lys Glu Pro Ser Val Asn
705                 710                 715                 720

Thr Ala Leu Val Pro Gln Leu Ser Thr Ile Ser Arg Ala Leu Thr Pro
                725                 730                 735

Ser Pro Val Met Val Leu Glu Asn Ile Glu Pro Glu Ile Val Tyr Ala
            740                 745                 750

Gly Tyr Asp Ser Ser Lys Pro Asp Thr Ala Glu Asn Leu Leu Ser Thr
        755                 760                 765

Leu Asn Arg Leu Ala Gly Lys Gln Met Ile Gln Val Val Lys Trp Ala
    770                 775                 780

Lys Val Leu Pro Gly Phe Lys Asn Leu Pro Leu Glu Asp Gln Ile Thr
785                 790                 795                 800

Leu Ile Gln Tyr Ser Trp Met Cys Leu Ser Ser Phe Ala Leu Ser Trp
                805                 810                 815

Arg Ser Tyr Lys His Thr Asn Ser Gln Phe Leu Tyr Phe Ala Pro Asp
            820                 825                 830

Leu Val Phe Asn Glu Leu Leu Ala Arg Val Arg Glu Gly
        835                 840                 845
```

<210> 34
<211> 837
<212> PRT
<213> HUMAN

<400> 34

```
Met Glu Thr Lys Gly Tyr His Ser Leu Pro Glu Gly Leu Asp Met Glu
  1               5                  10                  15

Arg Arg Trp Gly Gln Val Ser Gln Ala Val Glu Arg Ser Ser Leu Gly
             20                  25                  30

Pro Thr Glu Arg Thr Asp Glu Asn Asn Tyr Met Glu Ile Val Asn Val
         35                  40                  45

Ser Cys Val Ser Gly Ala Ile Pro Asn Asn Ser Thr Gln Gly Ser Ser
     50                  55                  60

Lys Glu Lys Gln Glu Leu Leu Pro Cys Leu Gln Gln Asp Asn Asn Arg
 65                  70                  75                  80

Pro Gly Ile Leu Thr Ser Asp Ile Lys Thr Glu Leu Glu Ser Lys Glu
                 85                  90                  95

Leu Ser Ala Thr Val Ala Glu Ser Met Gly Leu Tyr Met Asp Ser Val
            100                 105                 110

Arg Asp Ala Asp Tyr Ser Tyr Glu Gln Gln Asn Gln Gln Gly Ser Met
        115                 120                 125

Ser Pro Ala Lys Ile Tyr Gln Asn Val Glu Gln Leu Val Lys Phe Tyr
    130                 135                 140

Lys Gly Asn Gly His Arg Pro Ser Thr Leu Ser Cys Val Asn Thr Pro
145                 150                 155                 160

Leu Arg Ser Phe Met Ser Asp Ser Gly Ser Ser Val Asn Gly Gly Val
                165                 170                 175

Met Arg Ala Ile Val Lys Ser Pro Ile Met Cys His Glu Lys Ser Pro
            180                 185                 190

Ser Val Cys Ser Pro Leu Asn Met Thr Ser Ser Val Cys Ser Pro Ala
        195                 200                 205

Gly Ile Asn Ser Val Ser Ser Thr Thr Ala Ser Phe Gly Ser Phe Pro
    210                 215                 220

Val His Ser Pro Ile Thr Gln Gly Thr Pro Leu Thr Cys Ser Pro Asn
225                 230                 235                 240

Ala Glu Asn Arg Gly Ser Arg Ser His Ser Pro Ala His Ala Ser Asn
                245                 250                 255

Val Gly Ser Pro Leu Ser Ser Pro Leu Ser Ser Met Lys Ser Ser Ile
            260                 265                 270

Ser Ser Pro Pro Ser His Cys Ser Val Lys Ser Pro Val Ser Ser Pro
        275                 280                 285

Asn Asn Val Thr Leu Arg Ser Ser Val Ser Ser Pro Ala Asn Ile Asn
    290                 295                 300

Asn Ser Arg Cys Ser Val Ser Ser Pro Ser Asn Thr Asn Asn Arg Ser
305                 310                 315                 320
```

```
Thr Leu Ser Ser Pro Ala Ala Ser Thr Val Gly Ser Ile Cys Ser Pro
                325                 330                 335

Val Asn Asn Ala Phe Ser Tyr Thr Ala Ser Gly Thr Ser Ala Gly Ser
            340                 345                 350

Ser Thr Leu Arg Asp Val Val Pro Ser Pro Asp Thr Gln Glu Lys Gly
        355                 360                 365

Ala Gln Glu Val Pro Phe Pro Lys Thr Glu Glu Val Glu Ser Ala Ile
    370                 375                 380

Ser Asn Gly Val Thr Gly Gln Leu Asn Ile Val Gln Tyr Ile Lys Pro
385                 390                 395                 400

Glu Pro Asp Gly Ala Phe Ser Ser Ser Cys Leu Gly Gly Asn Ser Lys
                405                 410                 415

Ile Asn Ser Asp Ser Ser Phe Ser Val Pro Ile Lys Gln Glu Ser Thr
            420                 425                 430

Lys His Ser Cys Ser Gly Thr Ser Phe Lys Gly Asn Pro Thr Val Asn
        435                 440                 445

Pro Phe Pro Phe Met Asp Gly Ser Tyr Phe Ser Phe Met Asp Asp Lys
    450                 455                 460

Asp Tyr Tyr Ser Leu Ser Gly Ile Leu Gly Pro Pro Val Pro Gly Phe
465                 470                 475                 480

Asp Gly Asn Cys Glu Gly Ser Gly Phe Pro Val Gly Ile Lys Gln Glu
                485                 490                 495

Pro Asp Asp Gly Ser Tyr Tyr Pro Glu Ala Ser Ile Pro Ser Ser Ala
            500                 505                 510

Ile Val Gly Val Asn Ser Gly Gly Gln Ser Phe His Tyr Arg Ile Gly
        515                 520                 525

Ala Gln Gly Thr Ile Ser Leu Ser Arg Ser Ala Arg Asp Gln Ser Phe
    530                 535                 540

Gln His Leu Ser Ser Phe Pro Pro Val Asn Thr Leu Val Glu Ser Trp
545                 550                 555                 560

Lys Ser His Gly Asp Leu Ser Ser Arg Arg Ser Asp Gly Tyr Pro Val
                565                 570                 575

Leu Glu Tyr Ile Pro Glu Asn Val Ser Ser Ser Thr Leu Arg Ser Val
            580                 585                 590

Ser Thr Gly Ser Ser Arg Pro Ser Lys Ile Cys Leu Val Cys Gly Asp
        595                 600                 605

Glu Ala Ser Gly Cys His Tyr Gly Val Val Thr Cys Gly Ser Cys Lys
    610                 615                 620

Val Phe Phe Lys Arg Ala Val Glu Gly Gln His Asn Tyr Leu Cys Ala
625                 630                 635                 640

Gly Arg Asn Asp Cys Ile Ile Asp Lys Ile Arg Arg Lys Asn Cys Pro
```

645 650 655

Ala Cys Arg Leu Gln Lys Cys Leu Gln Ala Gly Met Asn Leu Gly Ala
660 665 670

Arg Lys Ser Lys Lys Leu Gly Lys Leu Lys Gly Ile His Glu Glu Gln
675 680 685

Pro Gln Gln Gln Gln Pro Pro Pro Pro Pro Pro Pro Pro Gln Ser Pro
690 695 700

Glu Glu Gly Thr Thr Tyr Ile Ala Pro Ala Lys Glu Pro Ser Val Asn
705 710 715 720

Thr Ala Leu Val Pro Gln Leu Ser Thr Ile Ser Arg Ala Leu Thr Pro
725 730 735

Ser Pro Val Met Val Leu Glu Asn Ile Glu Pro Glu Ile Val Tyr Ala
740 745 750

Gly Tyr Asp Ser Ser Lys Pro Asp Thr Ala Glu Asn Leu Leu Ser Thr
755 760 765

Leu Asn Arg Leu Ala Gly Lys Gln Met Ile Gln Val Val Lys Trp Ala
770 775 780

Lys Val Leu Pro Gly Phe Lys Asn Leu Pro Leu Glu Asp Gln Ile Thr
785 790 795 800

Leu Ile Gln Tyr Ser Trp Met Cys Leu Ser Ser Phe Ala Leu Ser Trp
805 810 815

Arg Ser Tyr Lys His Thr Asn Ser Gln Phe Leu Tyr Phe Ala Pro Asp
820 825 830

Leu Val Phe Asn Glu
835

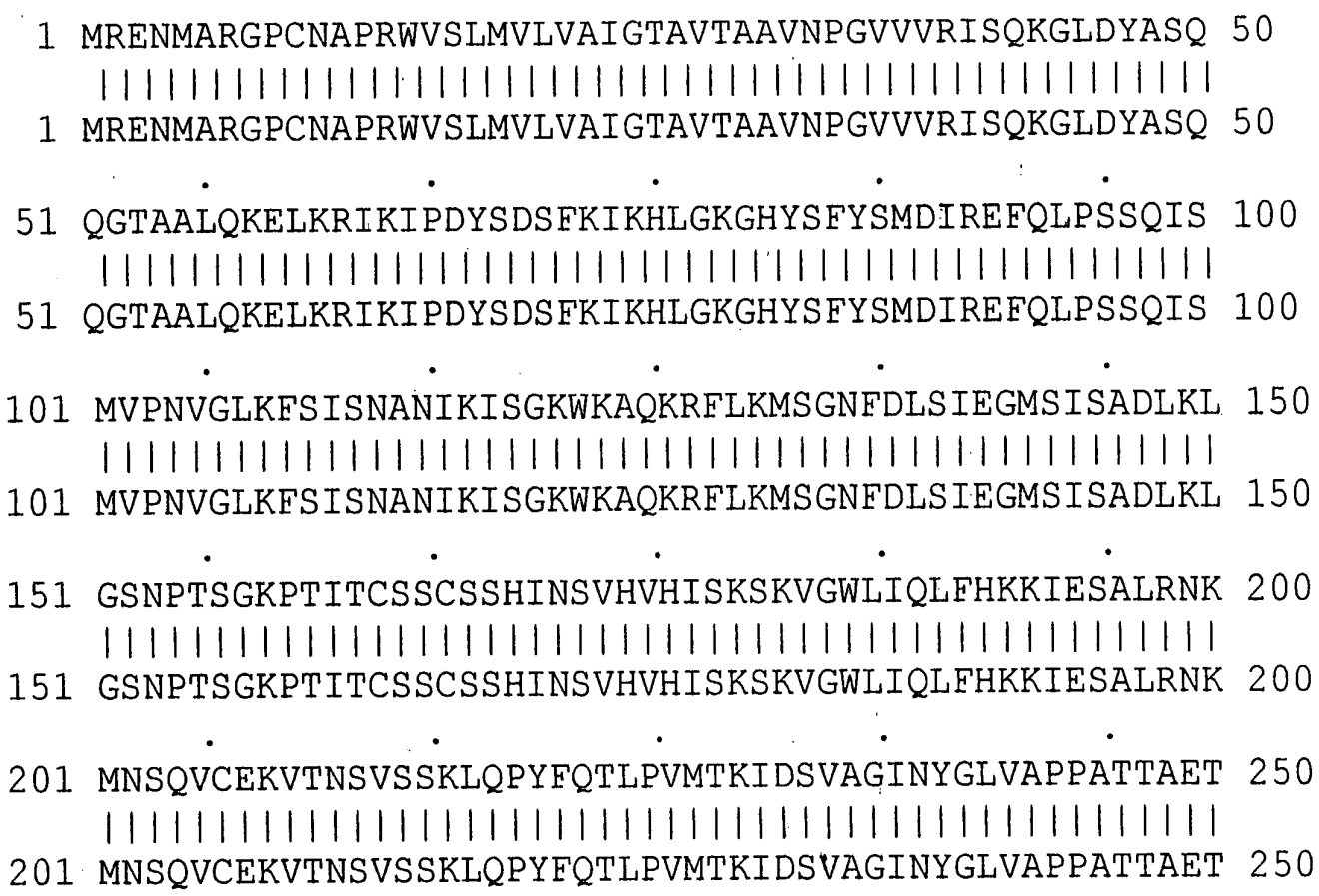

1 MRENMARGPCNAPRWVSLMVLVAIGTAVTAAVNPGVVVRISQKGLDYASQ 50
1 MRENMARGPCNAPRWVSLMVLVAIGTAVTAAVNPGVVVRISQKGLDYASQ 50
51 QGTAALQKELKRIKIPDYSDSFKIKHLGKGHYSFYSMDIREFQLPSSQIS 100
51 QGTAALQKELKRIKIPDYSDSFKIKHLGKGHYSFYSMDIREFQLPSSQIS 100
101 MVPNVGLKFSISNANIKISGKWKAQKRFLKMSGNFDLSIEGMSISADLKL 150
101 MVPNVGLKFSISNANIKISGKWKAQKRFLKMSGNFDLSIEGMSISADLKL 150
151 GSNPTSGKPTITCSSCSSHINSVHVHISKSKVGWLIQLFHKKIESALRNK 200
151 GSNPTSGKPTITCSSCSSHINSVHVHISKSKVGWLIQLFHKKIESALRNK 200
201 MNSQVCEKVTNSVSSKLQPYFQTLPVMTKIDSVAGINYGLVAPPATTAET 250
201 MNSQVCEKVTNSVSSKLQPYFQTLPVMTKIDSVAGINYGLVAPPATTAET 250

Fig. 1

```
251 LDVQMKGEFYSENHHNPPPFAPPVMEFPAAHDRMVYLGLSDYFFNTAGLV 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
251 LDVQMKGEFYSENHHNPPPFAPPVMEFPAAHDRMVYLGLSDYFFNTAGLV 300

301 YQEAGVLKMTLRDDMIPKESKFRLTTKFFGTFLPEVAKKFPNMKIQIHVS 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
301 YQEAGVLKMTLRDDMIPKESKFRLTTKFFGTFLPEVAKKFPNMKIQIHVS 350

351 ASTPPHLSVQPTGLTFYPAVDVQAFAVLPNSSLASLFLIGM 391
    |||||||||||||||||||||||||||||||||||||||||
351 ASTPPHLSVQPTGLTFYPAVDVQAFAVLPNSSLASLFLIGM 391
```

Fig. 1 (Cont.)

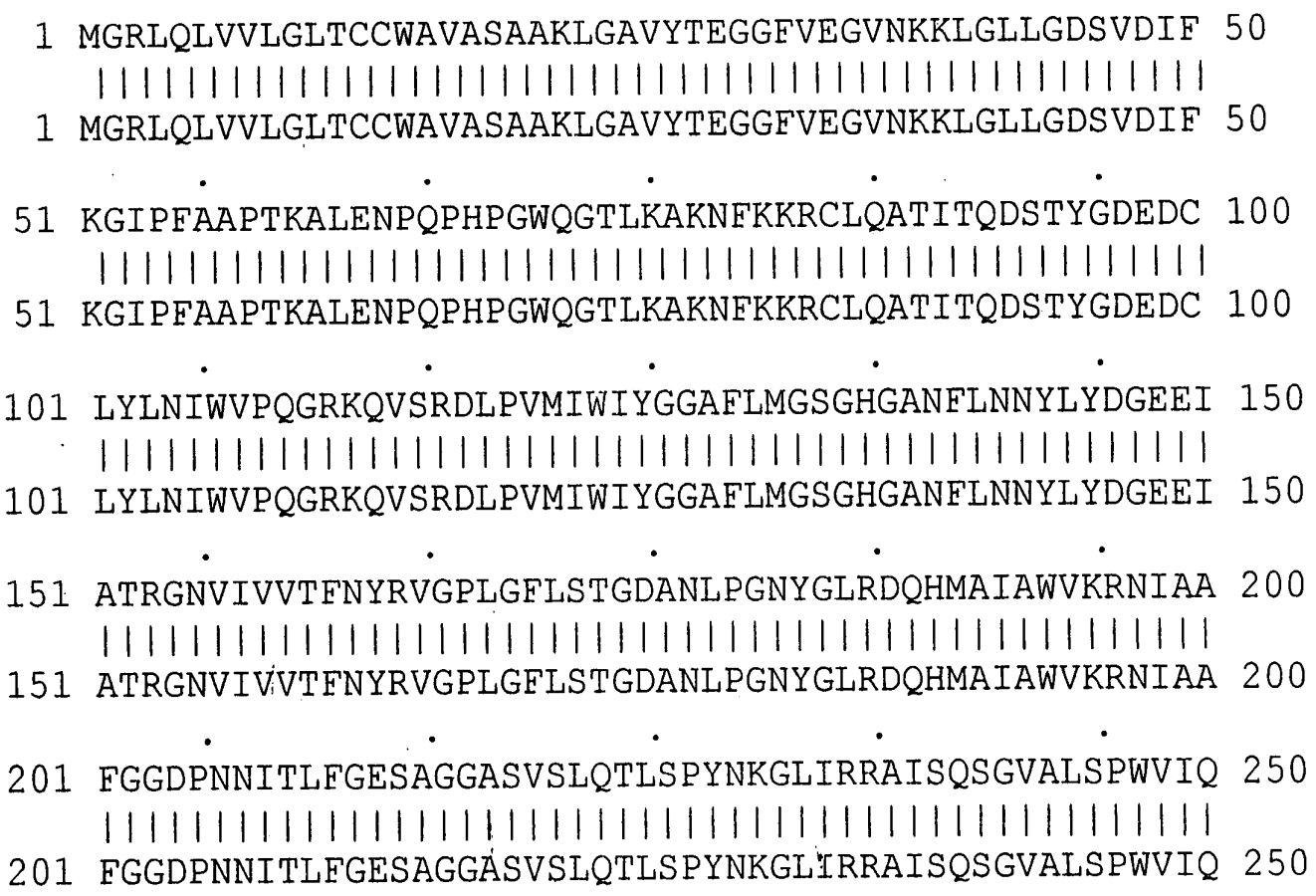

1 MGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIF 50
1 MGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLGDSVDIF 50
51 KGIPFAAPTKALENPQPHPGWQGTLKAKNFKKRCLQATITQDSTYGDEDC 100
51 KGIPFAAPTKALENPQPHPGWQGTLKAKNFKKRCLQATITQDSTYGDEDC 100
101 LYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEI 150
101 LYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFLNNYLYDGEEI 150
151 ATRGNVIVVTFNYRVGPLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAA 200
151 ATRGNVIVVTFNYRVGPLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAA 200
201 FGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQ 250
201 FGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQ 250

Fig. 2

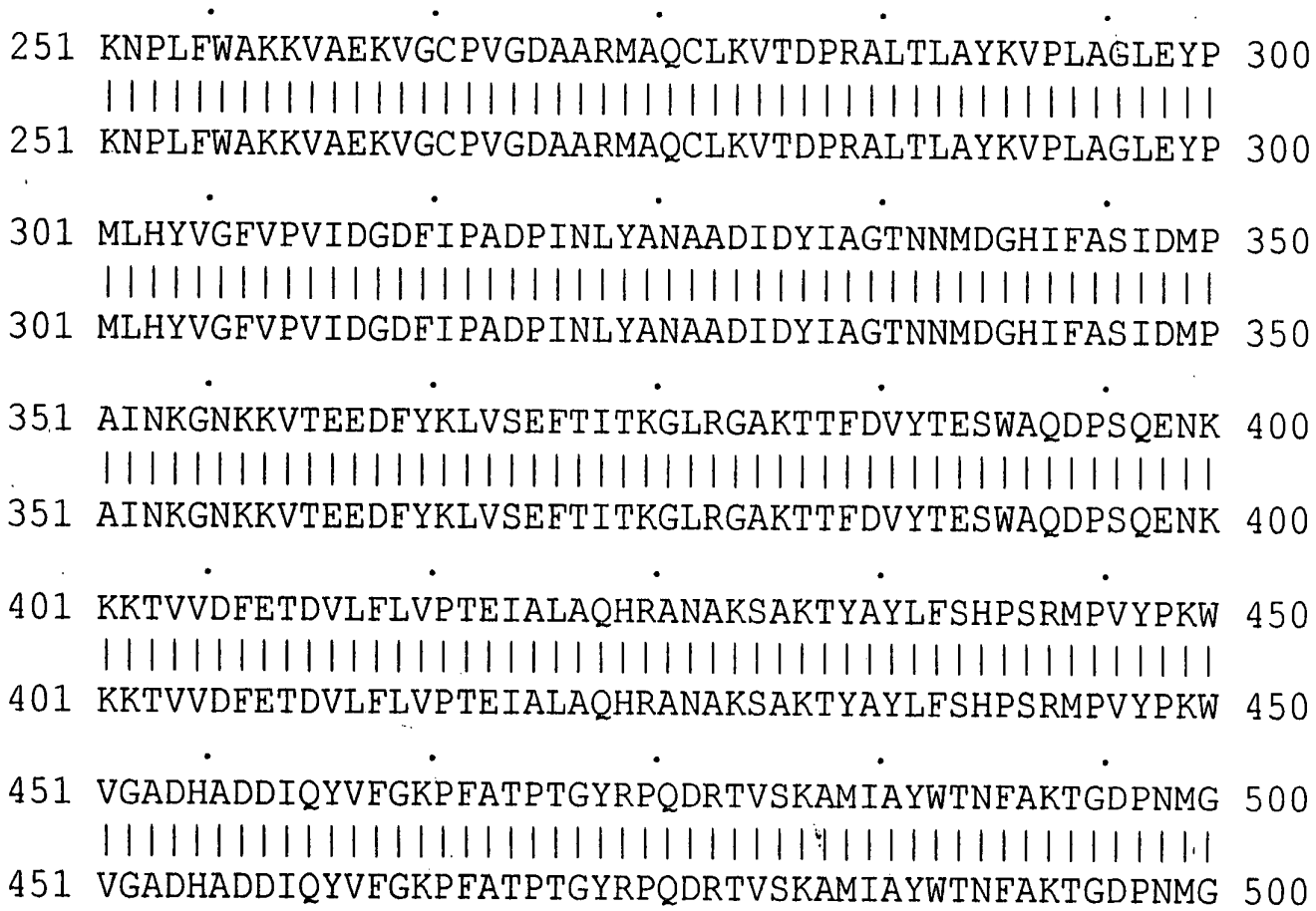

251 KNPLFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYP 300
251 KNPLFWAKKVAEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYP 300
301 MLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNNMDGHIFASIDMP 350
301 MLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNNMDGHIFASIDMP 350
351 AINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENK 400
351 AINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPSQENK 400
401 KKTVVDFETDVLFLVPTEIALAQHRANAKSAKTYAYLFSHPSRMPVYPKW 450
401 KKTVVDFETDVLFLVPTEIALAQHRANAKSAKTYAYLFSHPSRMPVYPKW 450
451 VGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMG 500
451 VGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMG 500

Fig. 2(Cont.)

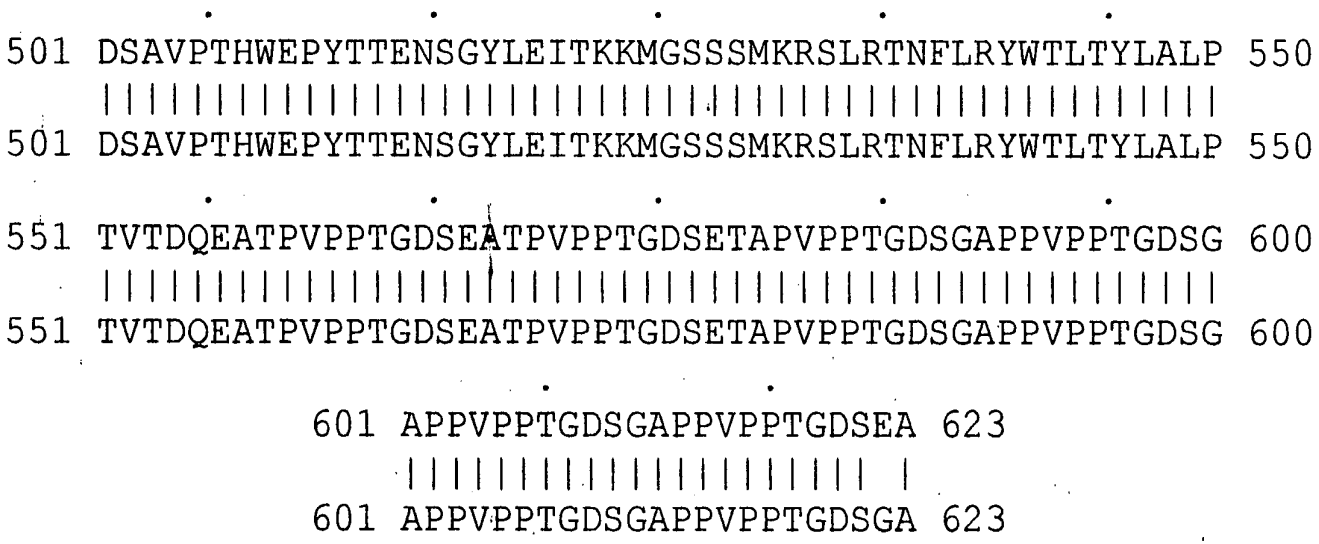

501 DSAVPTHWEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLTYLALP 550
501 DSAVPTHWEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLTYLALP 550
551 TVTDQEATPVPPTGDSEATPVPPTGDSETAPVPPTGDSGAPPVPPTGDSG 600
551 TVTDQEATPVPPTGDSEATPVPPTGDSETAPVPPTGDSGAPPVPPTGDSG 600
601 APPVPPTGDSGAPPVPPTGDSEA 623
601 APPVPPTGDSGAPPVPPTGDSGA 623

Fig. 2 (Cont.)

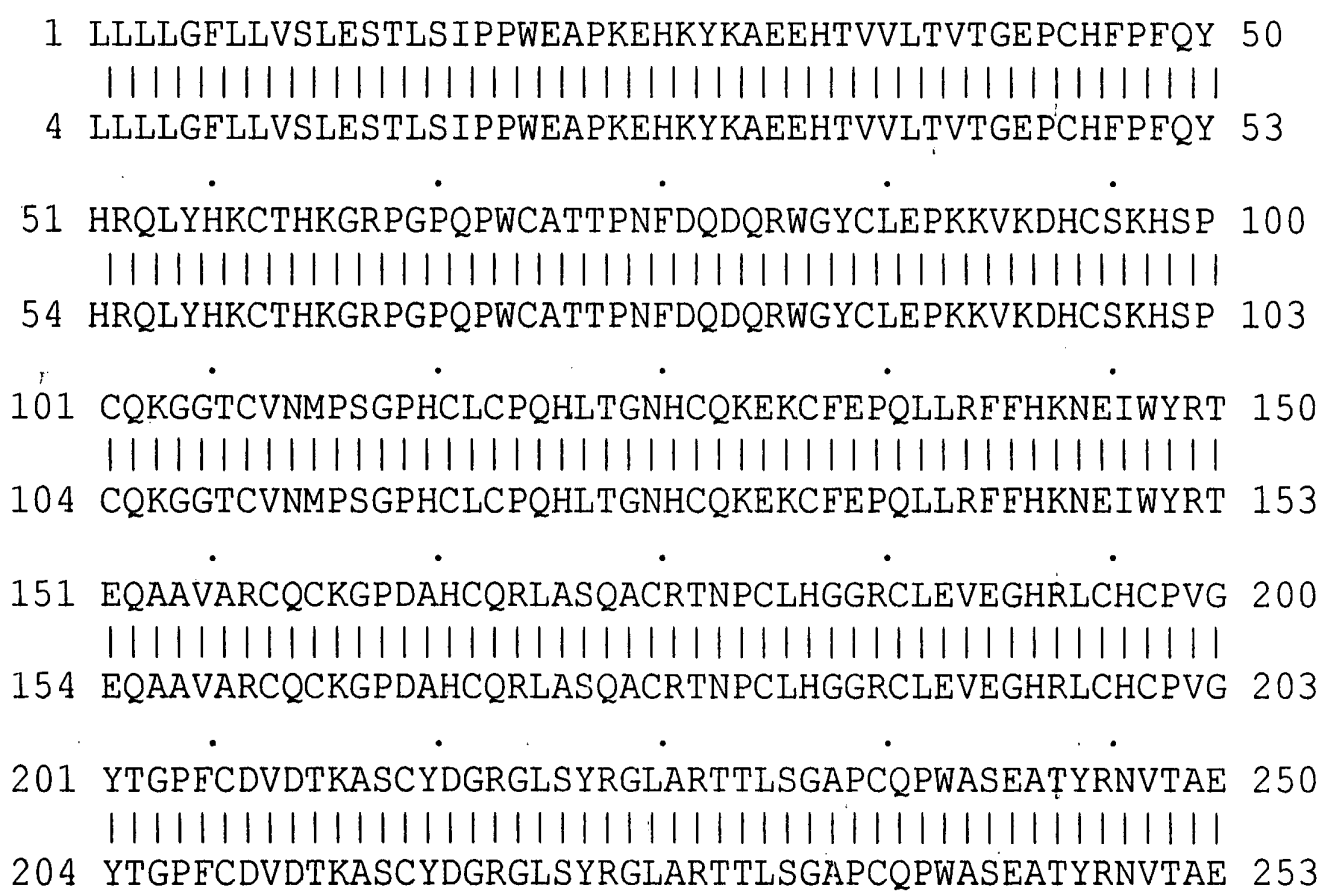

1 LLLLGFLLVSLESTLSIPPWEAPKEHKYKAEEHTVVLTVTGEPCHFPFQY 50
4 LLLLGFLLVSLESTLSIPPWEAPKEHKYKAEEHTVVLTVTGEPCHFPFQY 53
51 HRQLYHKCTHKGRPGPQPWCATTPNFDQDQRWGYCLEPKKVKDHCSKHSP 100
54 HRQLYHKCTHKGRPGPQPWCATTPNFDQDQRWGYCLEPKKVKDHCSKHSP 103
101 CQKGGTCVNMPSGPHCLCPQHLTGNHCQKEKCFEPQLLRFFHKNEIWYRT 150
104 CQKGGTCVNMPSGPHCLCPQHLTGNHCQKEKCFEPQLLRFFHKNEIWYRT 153
151 EQAAVARCQCKGPDAHCQRLASQACRTNPCLHGGRCLEVEGHRLCHCPVG 200
154 EQAAVARCQCKGPDAHCQRLASQACRTNPCLHGGRCLEVEGHRLCHCPVG 203
201 YTGPFCDVDTKASCYDGRGLSYRGLARTTLSGAPCQPWASEATYRNVTAE 250
204 YTGPFCDVDTKASCYDGRGLSYRGLARTTLSGAPCQPWASEATYRNVTAE 253

Fig. 3

```
251 QARNWGLGGHAFCRNPDNDIRPWCFVLNRDRLSWEYCDLAQCQTPTQAAP 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
254 QARNWGLGGHAFCRNPDNDIRPWCFVLNRDRLSWEYCDLAQCQTPTQAAP 303

301 PTPVSPRLHVPLMPAQPAPPKPQPTTRTPPQSQTPGALPAKREQPPSLTR 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
304 PTPVSPRLHVPLMPAQPAPPKPQPTTRTPPQSQTPGALPAKREQPPSLTR 353

351 NGPLSCGQRLRKSLSSMTRVVGGLVALRGAHPYIAALYWGHSFCAGSLIA 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
354 NGPLSCGQRLRKSLSSMTRVVGGLVALRGAHPYIAALYWGHSFCAGSLIA 403

401 PCWVLTAAHCLQDRPAPEDLTVVLGQERRNHSCEPCQTLAVRSYRLHEAF 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||
404 PCWVLTAAHCLQDRPAPEDLTVVLGQERRNHSCEPCQTLAVRSYRLHEAF 453

451 SPVSYQHDLALLRLQEDADGSCALLSPYVQPVCLPSGAARPSETTLCQVA 500
    ||||||||||||||||||||||||||||||||||||||||||||||||||
454 SPVSYQHDLALLRLQEDADGSCALLSPYVQPVCLPSGAARPSETTLCQVA 503
```

Fig. 3 (Cont.)

```
501 GWGHQFEGAEEYASFLQEAQVPFLSLERCSAPDVHGSSILPGMLCAGFLE 550
    ||||||||||||||||||||||||||||||||||||||||||||||||||
504 GWGHQFEGAEEYASFLQEAQVPFLSLERCSAPDVHGSSILPGMLCAGFLE 553

551 GGTDACAGELLAGWRPSPRPSAXSQVHSADCVFPTQGDSGGPLVCEDQAA 600
    ||||||                              ||||||||||||||||
554 GGTDAC..............................QGDSGGPLVCEDQAA 574

601 ERRLTLQGIISWGSGCGDRNKPGVYTDVAYYLAWIREHTVS 641
    |||||||||||||||||||||||||||||||||||||||||
575 ERRLTLQGIISWGSGCGDRNKPGVYTDVAYYLAWIREHTVS 615
```

Fig. 3 (Cont.)

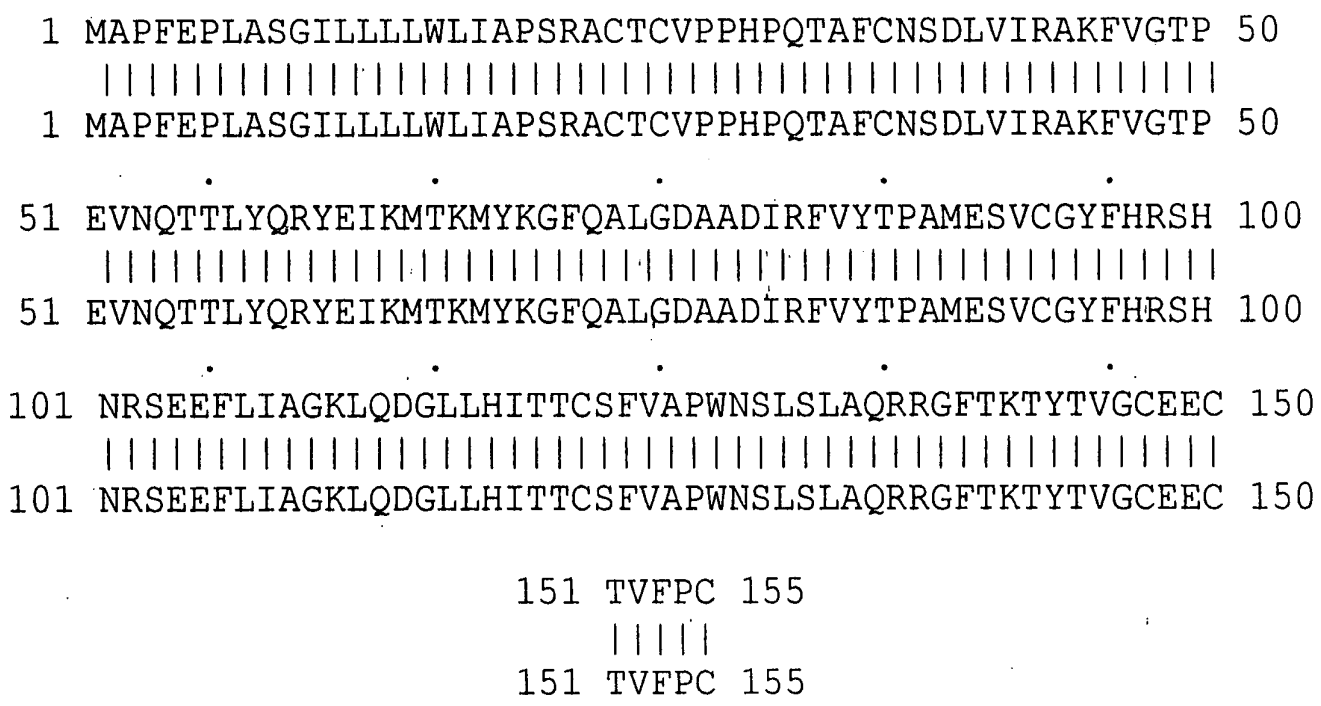

1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
101 NRSEEFLIAGKLQDGLLHITTCSFVAPWNSLSLAQRRGFTKTYTVGCEEC 150
101 NRSEEFLIAGKLQDGLLHITTCSFVAPWNSLSLAQRRGFTKTYTVGCEEC 150
151 TVFPC 155
151 TVFPC 155

Fig. 4

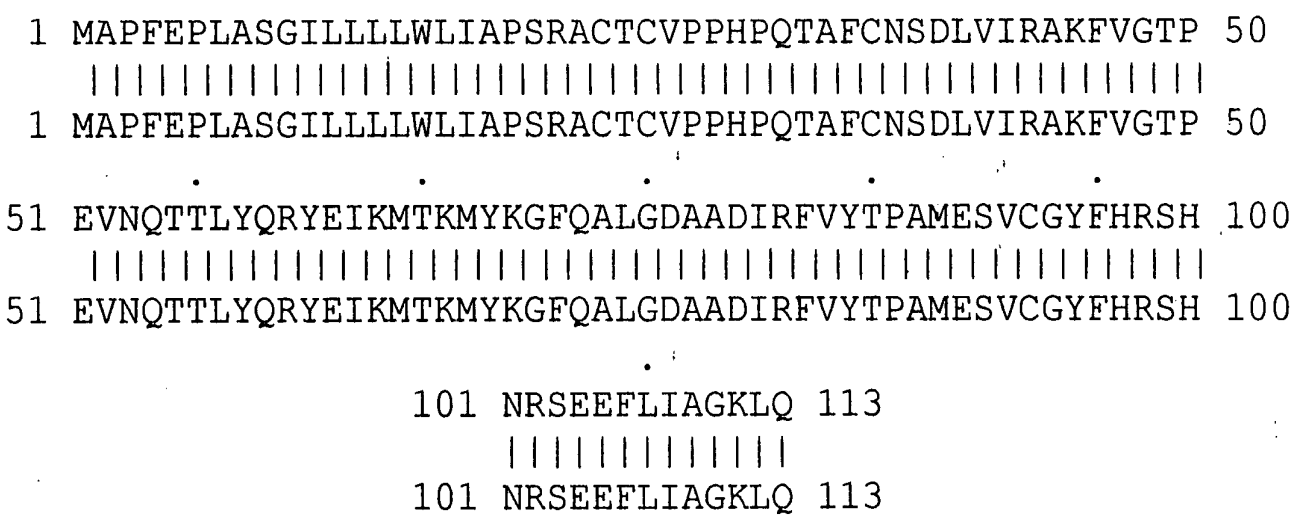

1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
101 NRSEEFLIAGKLQ 113
101 NRSEEFLIAGKLQ 113

Fig. 5

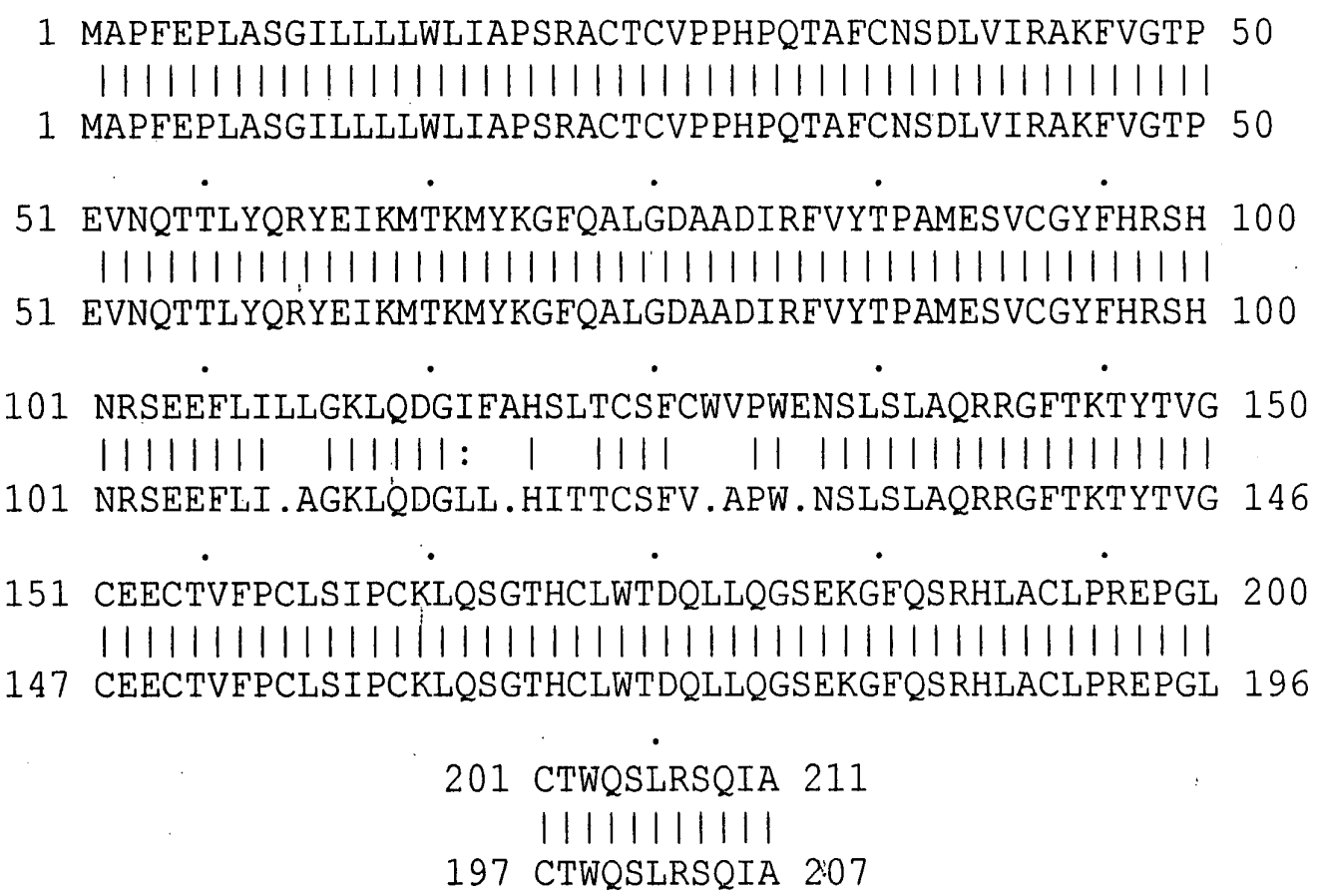

1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
101 NRSEEFLILLGKLQDGIFAHSLTCSFCWVPWENSLSLAQRRGFTKTYTVG 150
101 NRSEEFLI.AGKLQDGLL.HITTCSFV.APW.NSLSLAQRRGFTKTYTVG 146
151 CEECTVFPCLSIPCKLQSGTHCLWTDQLLQGSEKGFQSRHLACLPREPGL 200
147 CEECTVFPCLSIPCKLQSGTHCLWTDQLLQGSEKGFQSRHLACLPREPGL 196
201 CTWQSLRSQIA 211
197 CTWQSLRSQIA 207

Fig. 6

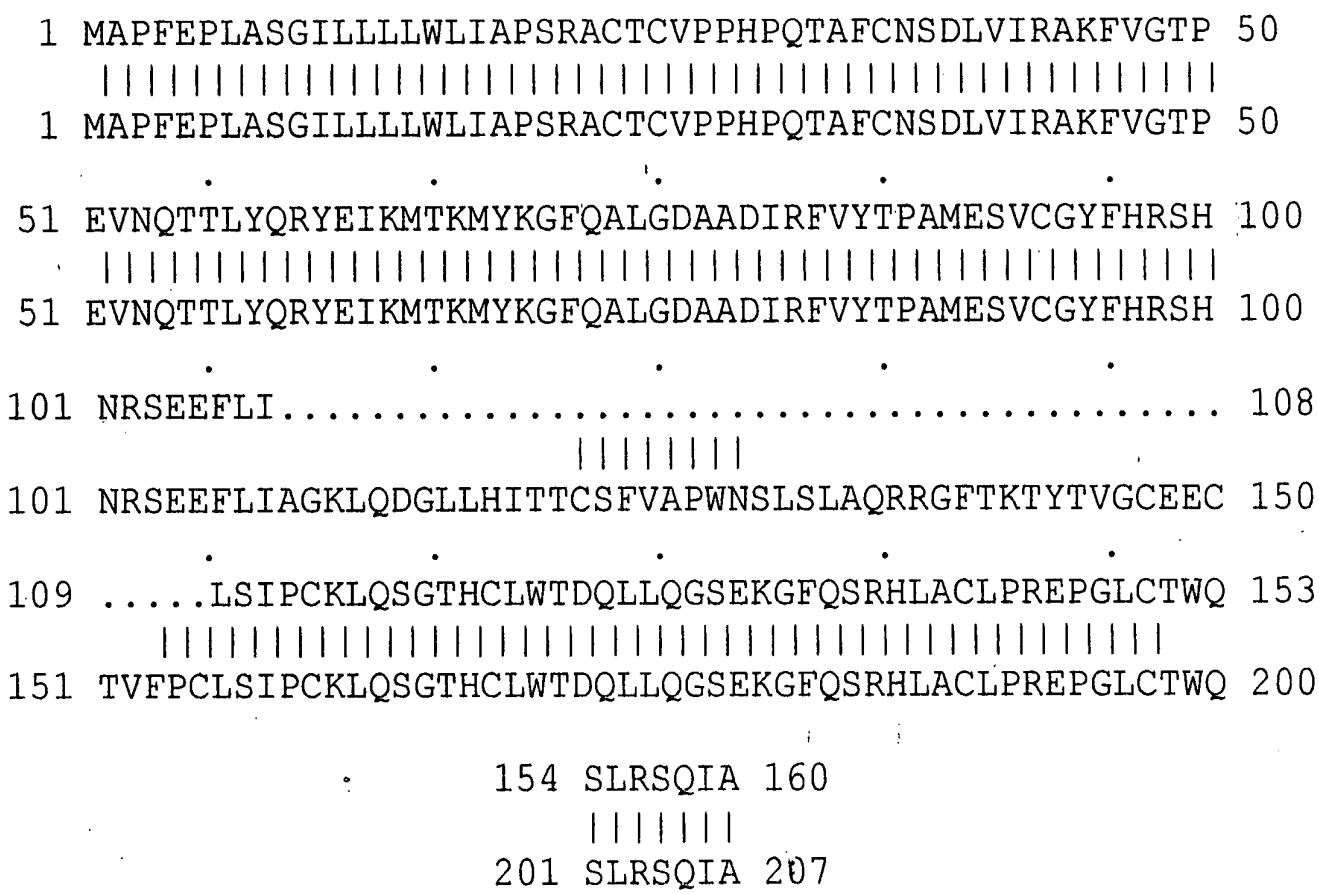

1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
101 NRSEEFLI..... 108
101 NRSEEFLIAGKLQDGLLHITTCSFVAPWNSLSLAQRRGFTKTYTVGCEEC 150
109 .....LSIPCKLQSGTHCLWTDQLLQGSEKGFQSRHLACLPREPGLCTWQ 153
151 TVFPCLSIPCKLQSGTHCLWTDQLLQGSEKGFQSRHLACLPREPGLCTWQ 200
154 SLRSQIA 160
201 SLRSQIA 207

Fig. 7

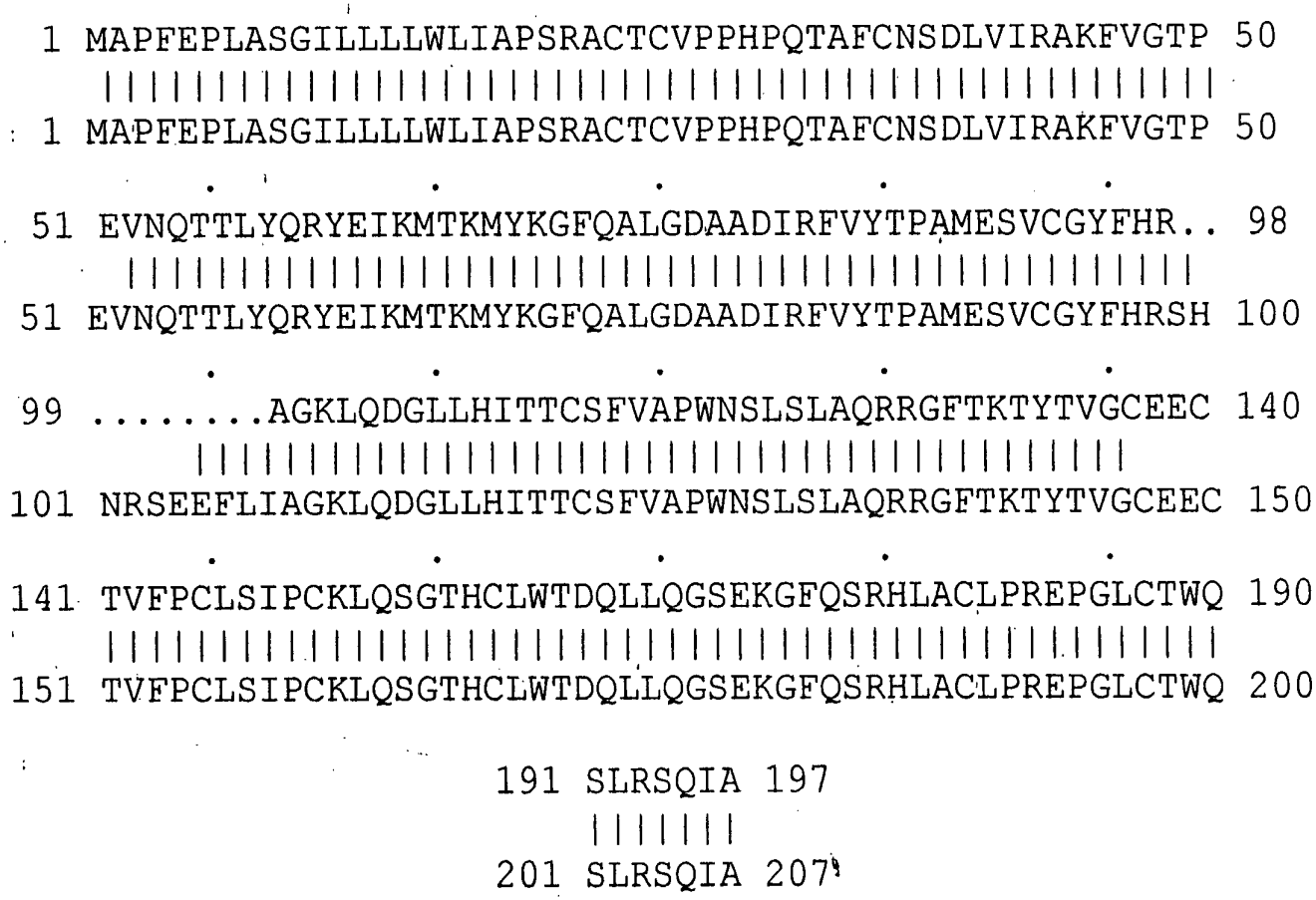

1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
1 MAPFEPLASGILLLLWLIAPSRACTCVPPHPQTAFCNSDLVIRAKFVGTP 50
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHR.. 98
51 EVNQTTLYQRYEIKMTKMYKGFQALGDAADIRFVYTPAMESVCGYFHRSH 100
99 .....AGKLQDGLLHITTCSFVAPWNSLSLAQRRGFTKTYTVGCEEC 140
101 NRSEEFLIAGKLQDGLLHITTCSFVAPWNSLSLAQRRGFTKTYTVGCEEC 150
141 TVFPCLSIPCKLQSGTHCLWTDQLLQGSEKGFQSRHLACLPREPGLCTWQ 190
151 TVFPCLSIPCKLQSGTHCLWTDQLLQGSEKGFQSRHLACLPREPGLCTWQ 200
191 SLRSQIA 197
201 SLRSQIA 207

Fig. 8

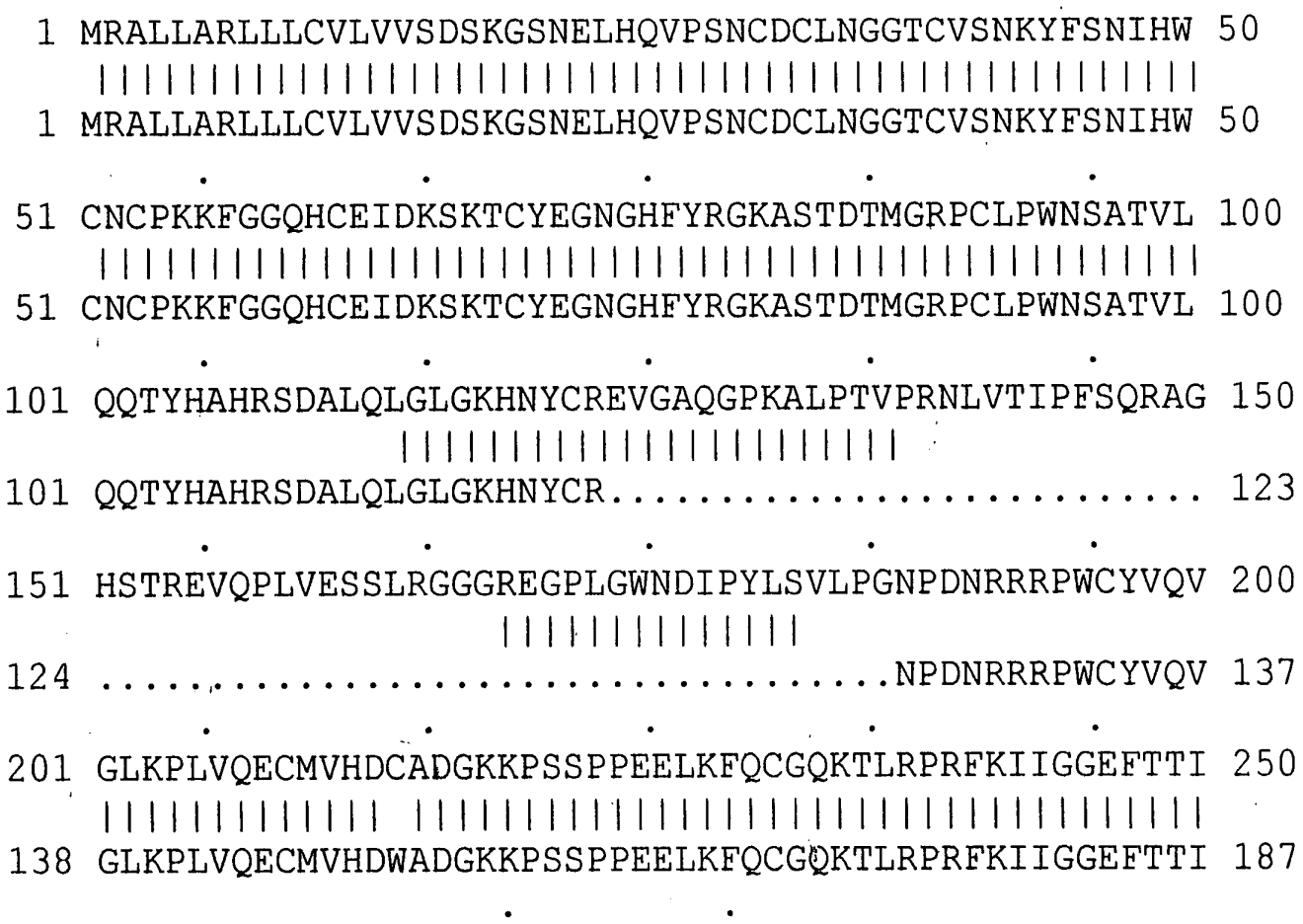

1 MRALLARLLLCVLVVSDSKGSNELHQVPSNCDCLNGGTCVSNKYFSNIHW 50
1 MRALLARLLLCVLVVSDSKGSNELHQVPSNCDCLNGGTCVSNKYFSNIHW 50
51 CNCPKKFGGQHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVL 100
51 CNCPKKFGGQHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVL 100
101 QQTYHAHRSDALQLGLGKHNYCREVGAQGPKALPTVPRNLVTIPFSQRAG 150
101 QQTYHAHRSDALQLGLGKHNYCR 123
151 HSTREVQPLVESSLRGGGREGPLGWNDIPYLSVLPGNPDNRRRPWCYVQV 200
124 NPDNRRRPWCYVQV 137
201 GLKPLVQECMVHDCADGKKPSSPPEELKFQCGQKTLRPRFKIIGGEFTTI 250
138 GLKPLVQECMVHDWADGKKPSSPPEELKFQCGQKTLRPRFKIIGGEFTTI 187

Fig. 9

```
251 ENQPWFAAIYRRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIV 300
    |||||||||||||||||||||||||||||||||||||||||||||||||||
188 ENQPWFAAIYRRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIV 237

301 YLGRSRLNSNTQGEMKFEVENLILHKDYSADTLAHHNDIALLKIRSKEGR 350
    |||||||||||||||||||||||||||||||||||||||||||||||||||
238 YLGRSRLNSNTQGEMKFEVENLILHKDYSADTLAHHNDIALLKIRSKEGR 287

351 CAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMTVV 400
    |||||||||||||||||||||||||||||||||||||||||||||||||||
288 CAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMTVV 337

401 KLISHRECQQPHYYGSEVTTKMLCAADPQWKTDSCQGDSGGPLVCSLQGR 450
    |||||||||||||||||||||||||||||||||||||||||||||||||| |
338 KLISHRECQQPHYYGSEVTTKMLCAADPQWKTDSCQGDSGGPLVCSLQCR 387

451 MTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKEENGLAL 494
    |||||||||||||||||||||||||||||||||||||||||||| |
388 MTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKEENGLVL 431
```

Fig. 9 (Cont.)

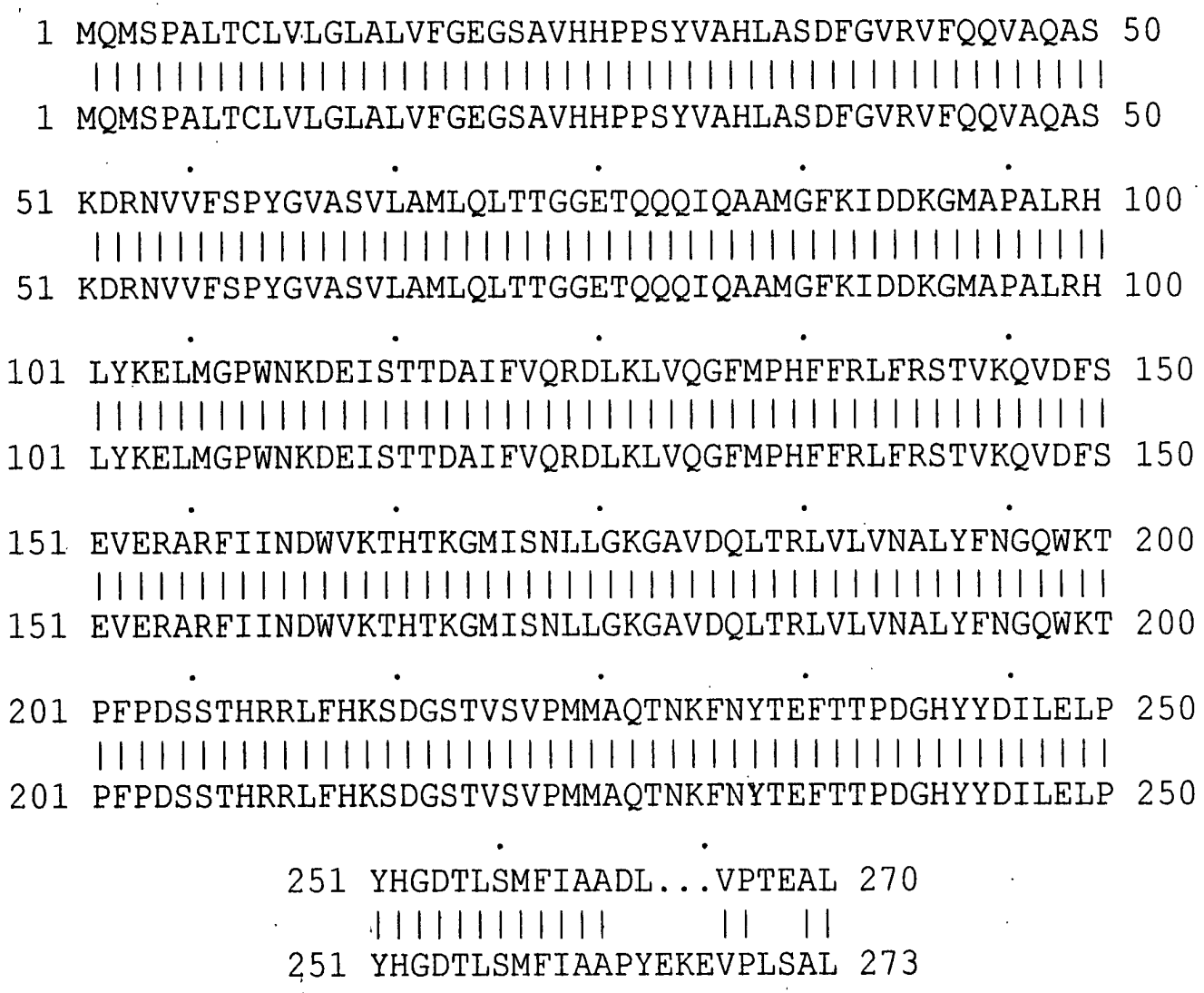

1 MQMSPALTCLVLGLALVFGEGSAVHHPPSYVAHLASDFGVRVFQQVAQAS 50
1 MQMSPALTCLVLGLALVFGEGSAVHHPPSYVAHLASDFGVRVFQQVAQAS 50
51 KDRNVVFSPYGVASVLAMLQLTTGGETQQQIQAAMGFKIDDKGMAPALRH 100
51 KDRNVVFSPYGVASVLAMLQLTTGGETQQQIQAAMGFKIDDKGMAPALRH 100
101 LYKELMGPWNKDEISTTDAIFVQRDLKLVQGFMPHFFRLFRSTVKQVDFS 150
101 LYKELMGPWNKDEISTTDAIFVQRDLKLVQGFMPHFFRLFRSTVKQVDFS 150
151 EVERARFIINDWVKTHTKGMISNLLGKGAVDQLTRLVLVNALYFNGQWKT 200
151 EVERARFIINDWVKTHTKGMISNLLGKGAVDQLTRLVLVNALYFNGQWKT 200
201 PFPDSSTHRRLFHKSDGSTVSVPMMAQTNKFNYTEFTTPDGHYYDILELP 250
201 PFPDSSTHRRLFHKSDGSTVSVPMMAQTNKFNYTEFTTPDGHYYDILELP 250
251 YHGDTLSMFIAADL...VPTEAL 270
251 YHGDTLSMFIAAPYEKEVPLSAL 273

Fig. 10

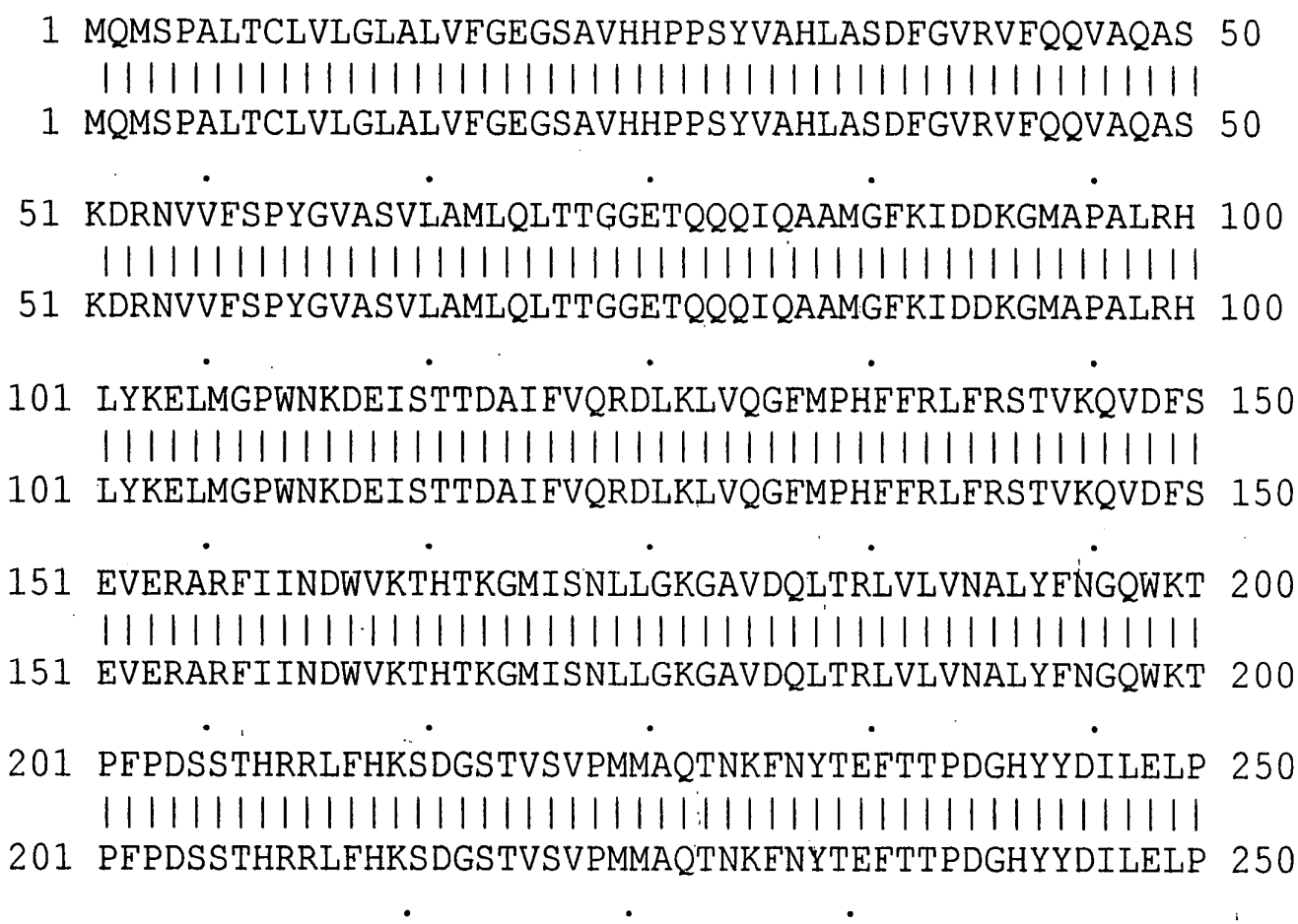

1 MQMSPALTCLVLGLALVFGEGSAVHHPPSYVAHLASDFGVRVFQQVAQAS 50
1 MQMSPALTCLVLGLALVFGEGSAVHHPPSYVAHLASDFGVRVFQQVAQAS 50
51 KDRNVVFSPYGVASVLAMLQLTTGGETQQQIQAAMGFKIDDKGMAPALRH 100
51 KDRNVVFSPYGVASVLAMLQLTTGGETQQQIQAAMGFKIDDKGMAPALRH 100
101 LYKELMGPWNKDEISTTDAIFVQRDLKLVQGFMPHFFRLFRSTVKQVDFS 150
101 LYKELMGPWNKDEISTTDAIFVQRDLKLVQGFMPHFFRLFRSTVKQVDFS 150
151 EVERARFIINDWVKTHTKGMISNLLGKGAVDQLTRLVLVNALYFNGQWKT 200
151 EVERARFIINDWVKTHTKGMISNLLGKGAVDQLTRLVLVNALYFNGQWKT 200
201 PFPDSSTHRRLFHKSDGSTVSVPMMAQTNKFNYTEFTTPDGHYYDILELP 250
201 PFPDSSTHRRLFHKSDGSTVSVPMMAQTNKFNYTEFTTPDGHYYDILELP 250

Fig. 11

```
251 YHGDTLSMFIAAPYEKEVPLSALTNILSAQLISHWKGNMTRLPRLLVLPK 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
251 YHGDTLSMFIAAPYEKEVPLSALTNILSAQLISHWKGNMTRLPRLLVLPK 300

301 FSLETEVDLRKPLENLGMTDMFRQFQADFTSLSDQEPLHVAQALQKVKIE 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
301 FSLETEVDLRKPLENLGMTDMFRQFQADFTSLSDQEPLHVAQALQKVKIE 350

351 VNESGTVASSSTAVIVSARMAPEEIIMDRPFLFVV 385
    ||||||||||||||||||||||||||||||||||||
351 VNESGTVASSSTAVIVSARMAPEEIIMDRPFLFVV 385
```

Fig. 11 (Cont.)


1 MTAASMGPVRVAFVVLLALCSRPAVGQNCSGPCRCPDEPAPRCPAGVSLV 50
1 MTAASMGPVRVAFVVLLALCSRPAVGQNCSGPCRCPDEPAPRCPAGVSLV 50
51 LDGCGCCRVCAKQLGELCTERDPCDPHKGLFCDFGSPANRKIGVCTAKDG 100
51 LDGCGCCRVCAKQLGELCTERDPCDPHKGLFCDFGSPANRKIGVCTAKDG 100
101 APCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSMDVRLPSPDCPL 150
101 APCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSMDVRLPSPDCPF 150
151 P.....LEDTFGPDPTMIRANCL 168
151 PRRVKLPGKCCEEWVCDEPKDQTVVGPALAAYRLEDTFGPDPTMIRANCL 200

Fig. 12

```
169 VQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCESDLEENI 218
    ||||||||||||||||||||||||||||||||||||||||||.||||||
201 VQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLEENI 250

219 KKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTT 268
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 KKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTT 300

269 LPVEFKCPDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 317
    |||||||||||||||||||||||||||||||||||||||||||||||||
301 LPVEFKCPDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349
```

Fig. 12 (Cont.)


3 MENSLRCVWVPKLAFVLFGASLLSAHLQVTGFQIKAFTALRFLSEPSDAV 52
1 MENSLRCVWVPKLAFVLFGASLLSAHLQVTGFQIKAFTALRFLSEPSDAV 50
53 TMRGGNVLLDCSAESDRGVPVIKWKKDAIHLALGMDERKQQLSNGSLLIQ 102
51 TMRGGNVLLDCSAESDRGVPVIKWKKDGIHLALGMDERKQQLSNGSLLIQ 100
103 NILHSRHHKPDEGLYQCEASLGDSGSIISRTAKVAVAGPLRFLSQTESVT 152
101 NILHSRHHKPDEGLYQCEASLGDSGSIISRTAKVAVAGPLRFLSQTESVT 150
153 AFMGDTVLLKCEVIGEPMPTIHWQKNQQDLTPIPGDSRVVVLPSGALQIS 202
151 AFMGDTVLLKCEVIGEPMPTIHWQKNQQDLTPIPGDSRVVVLPSGALQIS 200
203 RLQPGDIGIYRCSARNPASSRTGNEAEVRILSDPGLHRQLYFLQRPSNVV 252
201 RLQPGDIGIYRCSARNPASSRTGNEAEVRILSDPGLHRQLYFLQRPSNVV 250

Fig. 13

```
253 AIEGKDAVLECCVSGYPPPSFTWLRGEEVIQLRSKKYSLLGGSNLLISNV 302
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 AIEGKDAVLECCVSGYPPPSFTWLRGEEVIQLRSKKYSLLGGSNLLISNV 300

303 TDDDSGMYTCVVTYKNENISASAELTVLV 331
    ||||||||||||||||||||||||||||
301 TDDDSGMYTCVVTYKNENISASAELTVLV 329
```

Fig. 13 (Cont.)


3 MENSLRCVWVPKLAFVLFGASLLSAHLQVTGFQIKAFTALRFLSEPSDAV 52
1 MENSLRCVWVPKLAFVLFGASLLSAHLQVTGFQIKAFTALRFLSEPSDAV 50
53 TMRGGNVLLDCSAESDRGVPVIKWKKDAIHLALGMDERKQQLSNGSLLIQ 102
51 TMRGGNVLLDCSAESDRGVPVIKWKKDGIHLALGMDERKQQLSNGSLLIQ 100
103 NILHSRHHKPDEGLYQCEASLGDSGSIISRTAKVAVAGPLRFLSQTESVT 152
101 NILHSRHHKPDEGLYQCEASLGDSGSIISRTAKVAVAGPLRFLSQTESVT 150
153 AFMGDTVLLKCEVIGEPMPTIHWQKNQQDLTPIPGDSRVVVLPSGALQIS 202
151 AFMGDTVLLKCEVIGEPMPTIHWQKNQQDLTPIPGDSRVVVLPSGALQIS 200
203 RLQPGDIGIYRCSARNPASSRTGNEAEVRILSDPGLHRQLYFLQRPSNVV 252
201 RLQPGDIGIYRCSARNPASSRTGNEAEVRILSDPGLHRQLYFLQRPSNVV 250
253 AIEGKDAVLECCVSGYPPPSFTWLRGEEVIQLRSKKYSLLGGSNLLISNV 302
251 AIEGKDAVLECCVSGYPPPSFTWLRGEEVIQLRSKKYSLLGGSNLLISNV 300

Fig. 14

```
303 TDDDSGMYTCVVTYKNENISASAELTVLVPPWFLNHPSNLYAYESMDIEF 352
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
301 TDDDSGMYTCVVTYKNENISASAELTVLVPPWFLNHPSNLYAYESMDIEF 350

353 ECTVSGKPVPTVNWMKNGDVVIPSDYFQIVGGSNLRILGVVKSDEGFYQC 402
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
351 ECTVSGKPVPTVNWMKNGDVVIPSDYFQIVGGSNLRILGVVKSDEGFYQC 400

403 VAENEAGNAQTSAQLIVPKPAIPSSSVLPSAPRDVVPVLVSSRFVRLSWR 452
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
401 VAENEAGNAQTSAQLIVPKPAIPSSSVLPSAPRDVVPVLVSSRFVRLSWR 450

453 PPAEAKGNIQTFTVFFSREGDNRERALNTTQPGSLQLTVGNLKPEAMYTF 502
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
451 PPAEAKGNIQTFTVFFSREGDNRERALNTTQPGSLQLTVGNLKPEAMYTF 500

503 RVVAYNEWGPGESSQPIKVATQPELQVPGPVENLQAVSTSPTSILITWEP 552
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
501 RVVAYNEWGPGESSQPIKVATQPELQVPGPVENLQAVSTSPTSILITWEP 550

553 PAYANGPVQGYRLFCTEVSTGKEQNIEVDGLSYKLEGLKKFTEYSLRFLA 602
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
551 PAYANGPVQGYRLFCTEVSTGKEQNIEVDGLSYKLEGLKKFTEYSLRFLA 600
```

Fig. 14(Cont.)

```
603 YNRYGPGVSTDDITVVTLSDVPSAPPQNVSLEVVNSRSIKVSWLPPPSGT 652
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
601 YNRYGPGVSTDDITVVTLSDVPSAPPQNVSLEVVNSRSIKVSWLPPPSGT 650

653 QNGFITGYKIRHRKTTRRGEMETLEPNNLWYLFTGLEKGSQYSFQVSAMT 702
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
651 QNGFITGYKIRHRKTTRRGEMETLEPNNLWYLFTGLEKGSQYSFQVSAMT 700

703 VNGTGPPSNWYTAETPENDLDESQVPDQPSSLHVRPQTNCIIMSWTPPLN 752
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
701 VNGTGPPSNWYTAETPENDLDESQVPDQPSSLHVRPQTNCIIMSWTPPLN 750

753 PNIVVRGYIIGYGVGSPYAETVRVDSKQRYYSIERLESSSHYVISLKAFN 802
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
751 PNIVVRGYIIGYGVGSPYAETVRVDSKQRYYSIERLESSSHYVISLKAFN 800

803 NAGEGVPLYESATTRSITDPTDPVDYYPLLDDFPTSVPDLSTPMLPPVGV 852
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
801 NAGEGVPLYESATTRSITDPTDPVDYYPLLDDFPTSVPDLSTPMLPPVGV 850

853 QAVALTHDAVRVSWADNSVPKNQKTSEVRLYTVRWRTSFSASAKYKSEDT 902
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
851 QAVALTHDAVRVSWADNSVPKNQKTSEVRLYTVRWRTSFSASAKYKSEDT 900
```

Fig. 14(Cont.)

```
 903 TSLSYTATGLKPNTMYEFSVMVTKNRRSSTWSMTAHATTYEAAPTSAPKD 952
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
 901 TSLSYTATGLKPNTMYEFSVMVTKNRRSSTWSMTAHATTYEAAPTSAPKD 950

 953 FTVITREGKPRAVIVSWQPPLEANGKITAYILFYTLDKNIPIDDWIMETI 1002
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
 951 FTVITREGKPRAVIVSWQPPLEANGKITAYILFYTLDKNIPIDDWIMETI 1000

1003 SGDRLTHQIMDLNLDTMYYFRIQARNSKGVGPLSDPILFRTLKVEHPDKM 1052
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
1001 SGDRLTHQIMDLNLDTMYYFRIQARNSKGVGPLSDPILFRTLKVEHPDKM 1050

1053 ANDQGRHGDGGYWPVDTNLIDRSTLNEPPIGQMHPPHGSVTPQKNSNLLV 1102
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
1051 ANDQGRHGDGGYWPVDTNLIDRSTLNEPPIGQMHPPHGSVTPQKNSNLLV 1100

1103 IIVVTVGVITVLVVVIVAVICTRRSSAQQRKKRATHSAGKRKGSQKDLRP 1152
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
1101 IIVVTVGVITVLVVVIVAVICTRRSSAQQRKKRATHSAGKRKGSQKDLRP 1150

1153 PDLWIHHEEMEMKNIEKPSGTDPAGRDSPIQSCQDLTPVSHSQSETQLGS 1202
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
1151 PDLWIHHEEMEMKNIEKPSGTDPAGRDSPIQSCQDLTPVSHSQSETQLGS 1200
```

Fig. 14(Cont.)


1203 KSTSHSGQDTEEAGSSMSTLERSLAARRAPRAKLMIPMDAQSNNPAVVSA 1252
1201 KSTSHSGQDTEEAGSSMSTLERSLAARRAPRAKLMIPMDAQSNNPAVVSA 1250
1253 IPVPTLESAQYPGILPSPTCGYPHPQFTLRPVPFPTLSVDRGFGAGRSQS 1302
1251 IPVPTLESAQYPGILPSPTCGYPHPQFTLRPVPFPTLSVDRGFGAGRSQS 1300
1303 VSEGPTTQQPPMLPPSQPEHSSSEEAPSRTIPTACVRPTHPLRSFANPLL 1352
1301 VSEGPTTQQPPMLPPSQPEHSSSEEAPSRTIPTACVRPTHPLRSFANPLL 1350
1353 PPPMSAIEPKVPYTPLLSQPGPTLPKTHVKTASLGLAGKARSPLLPVSVP 1402
1351 PPPMSAIEPKVPYTPLLSQPGPTLPKTHVKTASLGLAGKARSPLLPVSVP 1400
1403 TAPEVSEESHKPTEDSANV 1421
1401 TAPEVSEESHKPTEDSANV 1419

Fig. 14(Cont.)

```
  1 MPGKRGLGWWWARLPLCLLLSLYGPWMPSSLGKPKGHPHMNSIRIDGDIT 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 MPGKRGLGWWWARLPLCLLLSLYGPWMPSSLGKPKGHPHMNSIRIDGDIT 50

 51 LGGLFPVHGRGSEGKPCGELKKEKGIHRLEAMLFALDRINNDPDLLPNIT 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 51 LGGLFPVHGRGSEGKPCGELKKEKGIHRLEAMLFALDRINNDPDLLPNIT 100

101 LGARILDTCSRDTHALEQSLTFVQALIEKDGTEVRCGSGGPPIITKPERV 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 LGARILDTCSRDTHALEQSLTFVQALIEKDGTEVRCGSGGPPIITKPERV 150

151 VGVIGASGSSVSIMVANILRLFKIPQISYASTAPDLSDNSRYDFFSRVVP 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 VGVIGASGSSVSIMVANILRLFKIPQISYASTAPDLSDNSRYDFFSRVVP 200

201 SDTYQAQAMVDIVRALKWNYVSTVASEGSYGESGVEAFIQKSREDGGVCI 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 SDTYQAQAMVDIVRALKWNYVSTVASEGSYGESGVEAFIQKSREDGGVCI 250

251 AQSVKIPREPKAGEFDKIIRRLLETSNARAVIIFANEDDIRRVLEAARRA 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 AQSVKIPREPKAGEFDKIIRRLLETSNARAVIIFANEDDIRRVLEAARRA 300
```

Fig. 15

```
301 NQTGHFFWMGSDSWGSKIAPVLHLEEVAEGAVTILPKRMSVR..... 342
    ||||||||||||||||||||||||||||||||||||||||||||
301 NQTGHFFWMGSDSWGSKIAPVLHLEEVAEGAVTILPKRMSVRGFDRYFSS 350
    .         .         .         .         .
343 .....DRERIGQDSAY 353
                .||||||||||
351 RTLDNNRRNIWFAEFWEDNFHCKLSRHALKKGSHVKKCTNRERIGQDSAY 400
    .         .         .         .         .
354 EQEGKVQFVIDAVYAMGHALHAMHRDLCPGRVGLCPRMDPVDGTQLLKYI 403
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 EQEGKVQFVIDAVYAMGHALHAMHRDLCPGRVGLCPRMDPVDGTQLLKYI 450
    .         .         .         .         .
404 RNVNFSGIAGNPVTFNENGDAPGRYDIYQYQLRNDSAEYKVIGSWTDHLH 453
    ||||||||||||||||||||||||||||||||||||||||||||||||||
451 RNVNFSGIAGNPVTFNENGDAPGRYDIYQYQLRNDSAEYKVIGSWTDHLH 500
    .         .         .         .         .
454 LRIERMHWPGSGQQLPRSICSLPCQPGERKKTVKGMPCCWHCEPCTGYQY 503
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 LRIERMHWPGSGQQLPRSICSLPCQPGERKKTVKGMPCCWHCEPCTGYQY 550
    .         .         .         .         .
504 QVDRYTCKTCPYDMRPTENRTGCRPIPIIKLEWGSPWAVLPLFLAVVGIA 553
    ||||||||||||||||||||||||||||||||||||||||||||||||||
551 QVDRYTCKTCPYDMRPTENRTGCRPIPIIKLEWGSPWAVLPLFLAVVGIA 600
              .         .
```

Fig. 15(Cont.)

```
554 ATLFVVITFVRYNDTPIVKASGRELSYVLLAGIFLCYATTFLMIAEPDLG 603
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
601 ATLFVVITFVRYNDTPIVKASGRELSYVLLAGIFLCYATTFLMIAEPDLG 650

604 TCSLRRIFLGLGMSISYAALLTKTNRIYRIFEQGKRSVSAPRFISPASQL 653
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
651 TCSLRRIFLGLGMSISYAALLTKTNRIYRIFEQGKRSVSAPRFISPASQL 700

654 AITFSLISLQLLGICVWFVVDPSHSVVDFQDQRTLDPRFARGVLKCDISD 703
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
701 AITFSLISLQLLGICVWFVVDPSHSVVDFQDQRTLDPRFARGVLKCDISD 750

704 LSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLA 753
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
751 LSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLA 800
```

Fig. 15(Cont.)

```
754 FIPIFFGTSQSADKLYIQTTTLTVSVSLSASVSLGMLYMPKVYIILFHPE 803
    |||||||||||||||||||||||||||||||||||||||||||||||||||
801 FIPIFFGTSQSADKLYIQTTTLTVSVSLSASVSLGMLYMPKVYIILFHPE 850

804 QNVPKRKRSLKAVVTAATMSNKFTQKGNFRPNGEAKSELCENLEAPALAT 853
    |||||||||||||||||||||||||||||||||||||||||||||||||||
851 QNVPKRKRSLKAVVTAATMSNKFTQKGNFRPNGEAKSELCENLEAPALAT 900

854 KQTYVTYTNHAI 865
    ||||||||||||
901 KQTYVTYTNHAI 912
```

Fig. 15(Cont.)

```
  1 METKGYHSLPEGLDMERRWGQVSQAVERSSLGPTERTDENNYMEIVNVSC 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 METKGYHSLPEGLDMERRWGQVSQAVERSSLGPTERTDENNYMEIVNVSC 50

 51 VSGAIPNNSTQGSSKEKQELLPCLQQDNNRPGILTSDIKTELESKELSAT 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 51 VSGAIPNNSTQGSSKEKQELLPCLQQDNNRPGILTSDIKTELESKELSAT 100

101 VAESMGLYMDSVRDADYSYEQQNQQGSMSPAKIYQNVEQLVKFYKGNGHR 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 VAESMGLYMDSVRDADYSYEQQNQQGSMSPAKIYQNVEQLVKFYKGNGHR 150

151 PSTLSCVNTPLRSFMSDSGSSVNGGVMRAIVKSPIMCHEKSPSVCSPLNM 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 PSTLSCVNTPLRSFMSDSGSSVNGGVMRAIVKSPIMCHEKSPSVCSPLNM 200

201 TSSVCSPAGINSVSSTTASFGSFPVHSPITQGTPLTCSPNAENRGSRSHS 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 TSSVCSPAGINSVSSTTASFGSFPVHSPITQGTPLTCSPNAENRGSRSHS 250

251 PAHASNVGSPLSSPLSSMKSSISSPPSHCSVKSPVSSPNNVTLRSSVSSP 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 PAHASNVGSPLSSPLSSMKSSISSPPSHCSVKSPVSSPNNVTLRSSVSSP 300
```

Fig. 16

```
301 ANINNSRCSVSSPSNTNNRSTLSSPAASTVGSICSPVNNAFSYTASGTSA 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
301 ANINNSRCSVSSPSNTNNRSTLSSPAASTVGSICSPVNNAFSYTASGTSA 350

351 GSSTLRDVVPSPDTQEKGAQEVPFPKTEEVESAISNGVTGQLNIVQYIKP 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
351 GSSTLRDVVPSPDTQEKGAQEVPFPKTEEVESAISNGVTGQLNIVQYIKP 400

401 EPDGAFSSSCLGGNSKINSDSSFSVPIKQESTKHSCSGTSFKGNPTVNPF 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
401 EPDGAFSSSCLGGNSKINSDSSFSVPIKQESTKHSCSGTSFKGNPTVNPF 450

451 PFMDGSYFSFMDDKDYYSLSGILGPPVPGFDGNCEGSGFPVGIKQEPDDG 500
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
451 PFMDGSYFSFMDDKDYYSLSGILGPPVPGFDGNCEGSGFPVGIKQEPDDG 500

501 SYYPEASIPSSAIVGVNSGGQSFHYRIGAQGTISLSRSARDQSFQHLSSF 550
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
501 SYYPEASIPSSAIVGVNSGGQSFHYRIGAQGTISLSRSARDQSFQHLSSF 550

551 PPVNTLVESWKSHGDLSSRRSDGYPVLEYIPENVSSSTLRSVSTGSSRPS 600
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
551 PPVNTLVESWKSHGDLSSRRSDGYPVLEYIPENVSSSTLRSVSTGSSRPS 600
```

Fig. 16 (Cont.)

```
601 KICLVCGDEASGCHYGVVTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKI 650
    ||||||||||||||||||||||||||||||||||||||||||||||||||
601 KICLVCGDEASGCHYGVVTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKI 650

651 RRKNCPACRLQKCLQAGMNLGARKSKKLGKLKGIHEEQPQQQQPPPPPPP 700
    ||||||||||||||||||||||||||||||||||||||||||||||||||
651 RRKNCPACRLQKCLQAGMNLGARKSKKLGKLKGIHEEQPQQQQPPPPPPP 700

701 PQSPEEGTTYIAPAKEPSVNTALVPQLSTISRALTPSPVMVLENIEPEIV 750
    ||||||||||||||||||||||||||||||||||||||||||||||||||
701 PQSPEEGTTYIAPAKEPSVNTALVPQLSTISRALTPSPVMVLENIEPEIV 750

751 YAGYDSSKPDTAENLLSTLNRLAGKQMIQVVKWAKVLPGFKNLPLEDQIT 800
    ||||||||||||||||||||||||||||||||||||||||||||||||||
751 YAGYDSSKPDTAENLLSTLNRLAGKQMIQVVKWAKVLPGFKNLPLEDQIT 800

801 LIQYSWMCLSSFALSWRSYKHTNSQFLYFAPDLVFNE 837
    |||||||||||||||||||||||||||||||||||||
801 LIQYSWMCLSSFALSWRSYKHTNSQFLYFAPDLVFNE 837
```

Fig. 16 (Cont.)

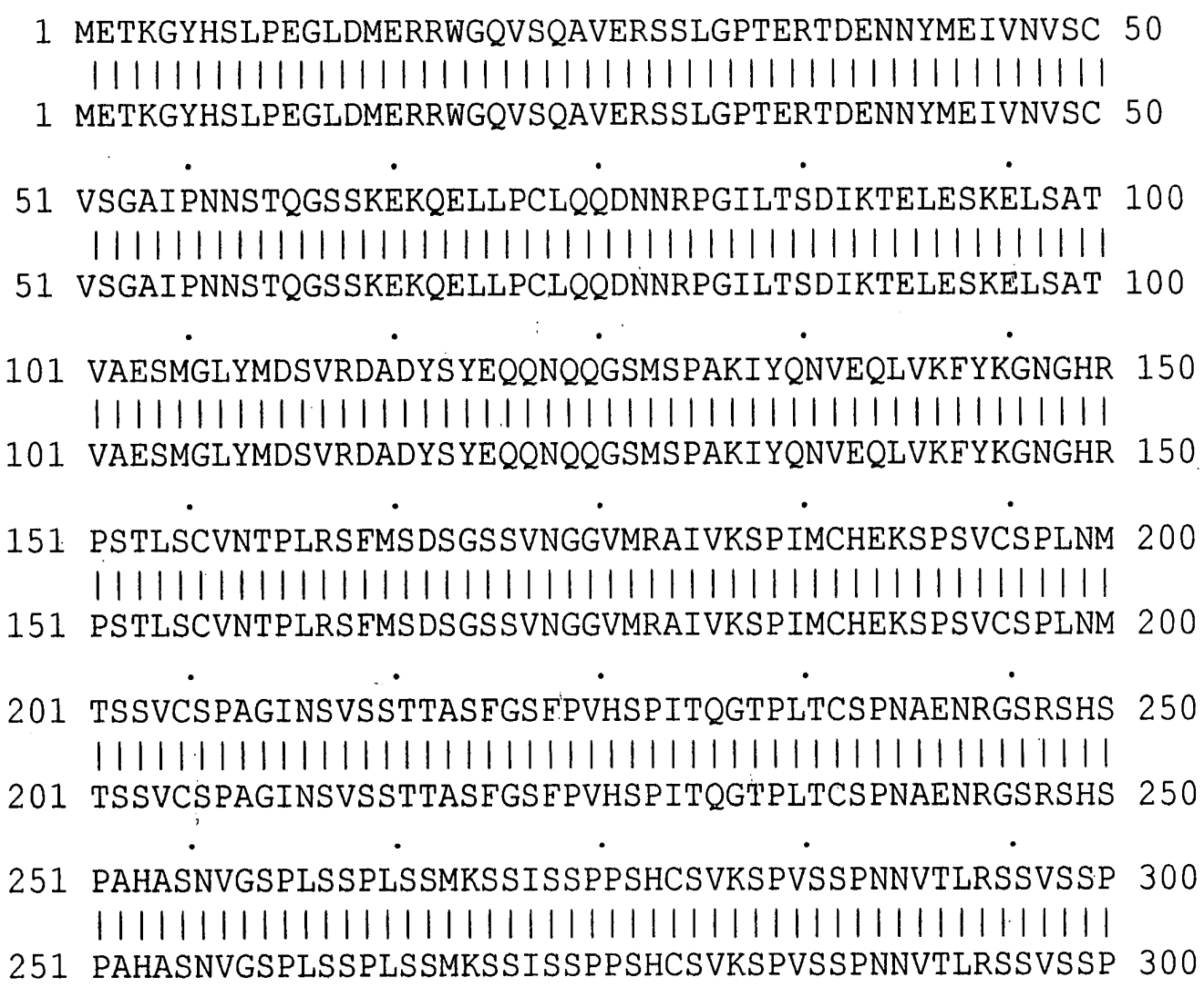

1 METKGYHSLPEGLDMERRWGQVSQAVERSSLGPTERTDENNYMEIVNVSC 50
1 METKGYHSLPEGLDMERRWGQVSQAVERSSLGPTERTDENNYMEIVNVSC 50
51 VSGAIPNNSTQGSSKEKQELLPCLQQDNNRPGILTSDIKTELESKELSAT 100
51 VSGAIPNNSTQGSSKEKQELLPCLQQDNNRPGILTSDIKTELESKELSAT 100
101 VAESMGLYMDSVRDADYSYEQQNQQGSMSPAKIYQNVEQLVKFYKGNGHR 150
101 VAESMGLYMDSVRDADYSYEQQNQQGSMSPAKIYQNVEQLVKFYKGNGHR 150
151 PSTLSCVNTPLRSFMSDSGSSVNGGVMRAIVKSPIMCHEKSPSVCSPLNM 200
151 PSTLSCVNTPLRSFMSDSGSSVNGGVMRAIVKSPIMCHEKSPSVCSPLNM 200
201 TSSVCSPAGINSVSSTTASFGSFPVHSPITQGTPLTCSPNAENRGSRSHS 250
201 TSSVCSPAGINSVSSTTASFGSFPVHSPITQGTPLTCSPNAENRGSRSHS 250
251 PAHASNVGSPLSSPLSSMKSSISSPPSHCSVKSPVSSPNNVTLRSSVSSP 300
251 PAHASNVGSPLSSPLSSMKSSISSPPSHCSVKSPVSSPNNVTLRSSVSSP 300

Fig. 17

```
301 ANINNSRCSVSSPSNTNNRSTLSSPAASTVGSICSPVNNAFSYTASGTSA 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 ANINNSRCSVSSPSNTNNRSTLSSPAASTVGSICSPVNNAFSYTASGTSA 350

351 GSSTLRDVVPSPDTQEKGAQEVPFPKTEEVESAISNGVTGQLNIVQYIKP 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 GSSTLRDVVPSPDTQEKGAQEVPFPKTEEVESAISNGVTGQLNIVQYIKP 400

401 EPDGAFSSSCLGGNSKINSDSSFSVPIKQESTKHSCSGTSFKGNPTVNPF 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 EPDGAFSSSCLGGNSKINSDSSFSVPIKQESTKHSCSGTSFKGNPTVNPF 450

451 PFMDGSYFSFMDDKDYYSLSGILGPPVPGFDGNCEGSGFPVGIKQEPDDG 500
    ||||||||||||||||||||||||||||||||||||||||||||||||||
451 PFMDGSYFSFMDDKDYYSLSGILGPPVPGFDGNCEGSGFPVGIKQEPDDG 500

501 SYYPEASIPSSAIVGVNSGGQSFHYRIGAQGTISLSRSARDQSFQHLSSF 550
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 SYYPEASIPSSAIVGVNSGGQSFHYRIGAQGTISLSRSARDQSFQHLSSF 550

551 PPVNTLVESWKSHGDLSSRRSDGYPVLEYIPENVSSSTLRSVSTGSSRPS 600
    ||||||||||||||||||||||||||||||||||||||||||||||||||
551 PPVNTLVESWKSHGDLSSRRSDGYPVLEYIPENVSSSTLRSVSTGSSRPS 600
```

Fig. 17(Cont.)

```
601 KICLVCGDEASGCHYGVVTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKI 650
    ||||||||||||||||||||||||||||||||||||||||||||||||||
601 KICLVCGDEASGCHYGVVTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKI 650

651 RRKNCPACRLQKCLQAGMNLGARKSKKLGKLKGIHEEQPQQQQPPPPPPP 700
    ||||||||||||||||||||||||||||||||||||||||||||||||||
651 RRKNCPACRLQKCLQAGMNLGARKSKKLGKLKGIHEEQPQQQQPPPPPPP 700

701 PQSPEEGTTYIAPAKEPSVNTALVPQLSTISRALTPSPVMVLENIEPEIV 750
    ||||||||||||||||||||||||||||||||||||||||||||||||||
701 PQSPEEGTTYIAPAKEPSVNTALVPQLSTISRALTPSPVMVLENIEPEIV 750

751 YAGYDSSKPDTAENLLSTLNRLAGKQMIQVVKWAKVLPGFKNLPLEDQIT 800
    ||||||||||||||||||||||||||||||||||||||||||||||||||
751 YAGYDSSKPDTAENLLSTLNRLAGKQMIQVVKWAKVLPGFKNLPLEDQIT 800

801 LIQYSWMCLSSFALSWRSYKHTNSQFLYFAPDLVFNE 837
    |||||||||||||||||||||||||||||||||||||
801 LIQYSWMCLSSFALSWRSYKHTNSQFLYFAPDLVFNE 837
```

Fig. 17 (Cont.)